DOCUMENT RESUME

ED 123 132 SE 020 896

AUTHOR	Cooney, Thomas J., Ed.; Bradbard, David A., Ed.
TITLE	Teaching Strategies: Papers from a Research Workshop.
INSTITUTION	ERIC Information Analysis Center for Science, Mathematics, and Environmental Education, Columbus, Ohio.; Georgia Univ., Athens. Georgia Center for the Study of Learning and Teaching Mathematics.
SPONS AGENCY	National Inst. of Education (DHEW), Washington, D.C.; National Science Foundation, Washington, D.C.
PUB DATE	May 76
NOTE	213p.
AVAILABLE FROM	Information Reference Center (ERIC/IRC), The Ohio State University, 1200 Chambers Road, 3rd Floor, Columbus, Ohio 43212 ($4.00)
EDRS PRICE	MF-$0.83 HC-$11.37 Plus Postage
DESCRIPTORS	Concept Teaching; Elementary Secondary Education; *Instruction; *Mathematics Education; *Research; *Research Reviews (Publications); Teacher Education; *Teaching Techniques

ABSTRACT

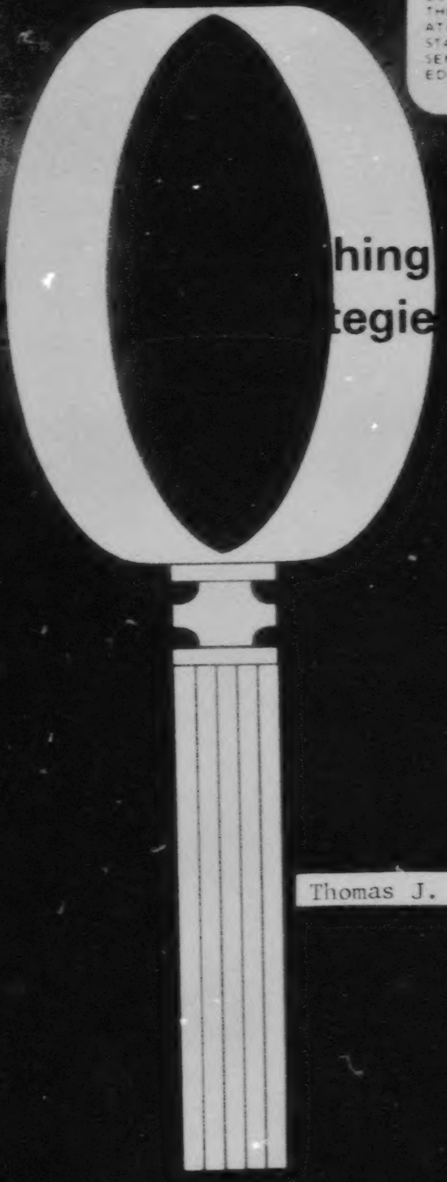
Nine papers presented at a research conference on strategies for teaching mathematics are presented in this volume. The first paper provides an overview of research on teaching strategies, defining a perspective on the subsequent papers. The second paper reviews the major strategies from a historical perspective. The third paper discusses the role of a theory in the development of teaching strategies. Four papers are concerned with research problems related to teaching strategies. The first of these deals with studies of efficacy of different strategies; the second concerns a comparison of teaching strategies which differed in the amount of information being taught and the amount of pupil-teacher interaction. More general research papers concern problems of designing studies of teaching strategies and a context for studying teaching strategies from a delivery-systems approach. The eighth paper discusses materials for teacher training. The final paper provides an integrative summary of research on teaching strategies. (SD)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED123132

U.S. DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY



hing
tegie

Thomas J. Cooney, Editor

2

SE 020 846

TEACHING

STRATEGIES

Papers from a Research Workshop

Sponsored by The Georgia Center
for the Study of Learning and
Teaching Mathematics
and the
Department of Mathematics Education
University of Georgia
Athens, Georgia

Thomas J. Cooney, Editor
David A. Bradbard, Technical Editor

ERIC Center for Science, Mathematics,
and Environmental Education
College of Education
The Ohio State University
1200 Chambers Road, Third Floor
Columbus, Ohio 43212

May, 1976

These papers were prepared as part of the activities of the Georgia Center for the Study of Learning and Teaching Mathematics, under Grant No. PES 7418491, National Science Foundation. The opinions expressed herein do not necessarily reflect the position or policy of the National Science Foundation.

This publication was prepared pursuant to a contract with the National Institute of Education, U.S. Department of Health, Education and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their judgment in professional and technical matters. Points of view or opinions do not, therefore, necessarily represent official National Institute of Education position or policy.

MATHEMATICS EDUCATION REPORTS

The Mathematics Education Reports series makes available recent analyses and syntheses of research and development efforts in mathematics education. We are pleased to make available as part of this series the papers from the Workshop on Teaching Strategies sponsored by the Georgia Center for the Study of Learning and Teaching Mathematics.

Other Mathematics Education Reports make available information concerning mathematics education documents analyzed at the ERIC Information Analysis Center for Science, Mathematics, and Environmental Education. These reports fall into three broad categories. Research reviews summarize and analyze recent research in specific areas of mathematics education. Resource guides identify and analyze materials and references for use by mathematics teachers at all levels. Special bibliographies announce the availability of documents and review the literature in selected interest areas of mathematics education. Reports in each of these categories may also be targeted for specific sub-populations of the mathematics education community.

Priorities for the development of future Mathematics Education Reports are established by the advisory board of the Center, in cooperation with the National Council of Teachers of Mathematics, the Special Interest Group for Research in Mathematics Education, and other professional groups in mathematics education. Individual comments on past Reports and suggestions for future Reports are always welcomed by the ERIC/SMEAC Center.

Jon L. Higgins
Associate Director

Contents

Acknowledgements and Overview
Leslie P. Steffe, Thomas J. Cooney, & Larry L. Hatfield vii

Research on Teaching Strategies: An Introduction
Thomas J. Cooney 1

Teaching Strategies: Historical and Contemporary Perspectives
B. Othanel Smith 23

Toward the Development of Pedagogical Theory in Mathematics
Kenneth B. Henderson 35

The Role of Relative Efficacy Studies in the Development of Mathematical Concept Teaching Strategies: Some Findings and Some Directions
John A. Dossey 51

An Empirical Comparison of Teaching Strategies Where the Amount of Content Information and Teacher-Pupil Interaction is Varied
Earl W. Swank 87

Design Problems in Research on Teaching Strategies in Mathematics
Richard L. Turner 115

A Research Context for Delivery Systems Research on Strategies for Teaching Mathematics
Kenneth A. Retzer 141

The Materials of Teacher Training
David Gliessman 171

Reflections from Research: Focusing on Teaching Strategies from Various Directions
Marilyn Suydam 185

Participants 207

Acknowledgements and Overview

The Georgia Center for the Study of Learning and Teaching Mathematics (GCSLTM) was started July 1, 1975, through a founding grant from the National Science Foundation. Various activities preceded the founding of the GCSLTM. The most significant was a conference held at Columbia University in October of 1970 on Piagetian Cognitive-Development and Mathematical Education. This conference was directed by the late Myron F. Rosskopf and jointly sponsored by the National Council of Teachers of Mathematics and the Department of Mathematical Education, Teachers College, Columbia University with a grant from the National Science Foundation. Following the October 1970 Conference, Professor Rosskopf spent the winter and spring quarters of 1971 as a visiting professor of Mathematics Education at the University of Georgia. During these two quarters, the editorial work was completed on the proceedings of the October conference and a Letter of Intent was filed in February of 1971 with the National Science Foundation to create a Center for Mathematical Education Research and Innovation. Professor Rosskopf's illness and untimely death made it impossible for him to develop the ideas contained in that Letter.

After much discussion among faculty in the Department of Mathematics Education at the University of Georgia, it was clear that a center devoted to the study of mathematics education ought to attack a broader range of problems than was stated in the Letter of Intent. As a result of these discussions, three areas of study were identified as being of primary interest in the initial year of the Georgia Center for the Study of Learning and Teaching Mathematics--Teaching Strategies, Concept Development, and Problem Solving. Thomas J. Cooney assumed directorship of the Teaching Strategies Project, Leslie P. Steffe the Concept Development Project, and Larry L. Hatfield the Problem Solving Project.

The GCSLTM is intended to be a long-term operation with the broad goal of improving mathematics education in elementary and secondary schools. To be effective, it was felt that the Center would have to include mathematics educators with interests commensurate with those of the project areas. Alternative organizational patterns were available--resident scholars, institutional consortia, or individual consortia. The latter organizational pattern was chosen because it was felt maximum participation would be then possible. In order to operationalize a concept of a consortia of individuals, five research workshops were held during the spring of 1975 at the University of Georgia. These workshops were (ordered by dates held) Teaching Strategies, Number and Measurement Concepts, Space and Geometry Concepts, Models for Learning Mathematics, and Problem Solving. Papers were commissioned for each workshop. It was necessary to commission papers for two reasons. First, current analyses and syntheses of the knowledge in the particular areas chosen for investigation were needed. Second, a catalyst for further research and development activities was needed--major problems had to be identified in the project areas on which work was needed.

Twelve working groups emerged from these workshops; three in Teaching Strategies, five in Concept Development, and four in Problem Solving. The three working groups in Teaching Strategies are: Differential Effects of Varying Teaching Strategies, John Dossey, Coordinator; Development of Protocol Materials to Depict Moves and Strategies, Kenneth Retzer, Coordinator; and Investigation of Certain Teacher Behavior That May Be Associated with Effective Teaching, Thomas J. Cooney, Coordinator. The five working groups in Concept Development are: Measurement Concepts, Thomas Romberg, Coordinator; Rational Number Concepts, Thomas Kieren, Coordinator; Cardinal and Ordinal Number Concepts, Leslie P. Steffe, Coordinator; Space and Geometry Concepts, Richard Lesh, Coordinator; and Models for Learning Mathematics, William Geeslin, Coordinator. The four working groups in Problem Solving are: Instruction in the Use of Key Organizers (Single Heuristics), Frank Lester, Coordinator; Instruction Organized to use Heuristics in Combinations, Phillip Smith, Coordinator; Instruction in Problem Solving Strategies, Douglas Grouws, Coordinator; and Task Variables for Problem Solving Research, Gerald Kulm, Coordinator. The twelve working groups are working as units somewhat independently of one another. As research and development emerges from working groups, it is envisioned that some working groups will merge naturally.

The publication program of the Center is of central importance to Center activities. Research and development monographs and school monographs will be issued, when appropriate, by each working group. The school monographs will be written in nontechnical language and are to be aimed at teacher educators and school personnel. Reports of single studies may be also published as technical reports.

All of the above plans and aspirations would not be possible if it were not for the existence of professional mathematics educators with the expertise in and commitment to research and development in mathematics education. The professional commitment of mathematics educators to the betterment of mathematics education in the schools has been vastly underestimated. In fact, the basic premise on which the GCSLTM is predicated is that there are a significant number of professional mathematics educators with a great deal of individual commitment to creative scholarship. There is no attempt on the part of the Center to buy this scholarship--only to stimulate it and provide a setting in which it can flourish.

The Center administration wishes to thank the individuals who wrote the excellent papers for the workshops, the participants who made the workshops possible, and the National Science Foundation for supporting financially the first year of Center operation. Various individuals have provided valuable assistance in preparing the papers given at the workshops for publication. Mr. David Bradbard provided technical editorship; Mrs. Julie Wetherbee, Mrs. Elizabeth Platt, Mrs. Kay Abney, and Mrs. Cheryl Hirstein, proved to be able typists; and Mr. Robert Petty drafted the figures. Mrs. Julie Wetherbee also provided expertise in the daily operation of the Center during its first year. One can only feel grateful for the existence of such capable and hardworking people.

Thomas J. Cooney
Director
Teaching Strategies

Leslie P. Steffe
Director
Concept Development
and
Director, GCSLTM

Larry L. Hatfield
Director
Problem Solving

Research on Teaching Strategies:

An Introduction

Thomas J. Cooney

University of Georgia

The Teaching Strategies Project has as its main objective the study of teaching and teacher training. In particular, the project is concerned with logical aspects of teacher behavior and the way in which these aspects may relate to effective teaching. The work of the Teaching Strategies Project originated in the theory of teaching advanced by Smith and Henderson. The papers in this collection focus on theory and research concerning the teaching of mathematics and possible directions for further research. In this introduction, the contents of the papers are highlighted and promising research on teaching mathematics that was done outside the project is identified and discussed.

Hyman (1971) described teaching as a triadic relationship, involving a teacher, at least one pupil, and the subject matter to be taught and learned. Henderson (1971) provided a more formalized conceptualization of teaching by interpreting teaching as a ternary relation $T(x, y, z)$. To Henderson, the domain of x constitutes "sequences of 'actions' of an object which, in terms of some criteria, is identified as a teacher" (p. 137). The domain of y is the set of teachable objects. The domain of z according to Henderson is "sequences of actions or behaviors of a person who, in terms of some set criteria, is identified as a learner" (p. 138). It is clear that the domain of z is a factor influencing the conscious actions of a teacher, as teachers behave differently with respect to the nature of the learner and the learner's respective actions. An artistic teacher alters his or her teaching strategies according to his or her perception of the status of the learner as evidenced by the learner's behaviors.

Although not as obvious, the domain of y is also a factor in determining the strategies a teacher utilizes. Henderson (1972) made the following observation with reference to the nature and influence on teaching of the domain of y:

> One can hypothesize that the kind of teachable object (value of y) makes a difference in teaching (value of x) just as the kind of behavior of the student (value of z) does. Surely a teacher should go about teaching an item of analytic knowledge or belief differently than he would an item of empirical knowledge or belief. And just as surely a teacher should draw a distinction between a factual statement and a value judgment and hence teach them differently. (p. 4)

To illustrate, consider three items of knowledge commonly found in secondary school mathematics.

1. Rhombus
2. A rhombus is a parallelogram with two adjacent sides congruent.
3. The diagonals of a rhombus are perpendicular.

Consider, now, what a teacher might do in teaching these items of knowledge. For item 1 it makes sense to point to a specific object that represents or does not represent a rhombus. But it does not make sense to point to a specific object and say that it does or does not represent items 2 or 3. This is because the referent of item 1 is a collection of objects whereas the other two items involve truth functional statements. But these two statements differ also. In 2 a common definition of a rhombus is given. As such, the statement is true by agreement or stipulation. In 3, however, the truth of the statement is predicated on the truth of other assertions--some definitional and some deduced from previously established statements. The nature of the teachable object in each case will determine, in part, how a teacher behaves and the nature of the teaching strategy selected.

It is generally recognized that teachers must have extensive knowledge of both subject matter and basic psychological principles applicable to classroom instruction. What is not as readily recognized is the relevance of epistemology in determining a teacher's behavior. As Henderson pointed out, the nature of the knowledge being taught also determines a teacher's actions. Smith (1969), in discussing this third kind of knowledge, wrote:

> It has only recently been recognized that there is another sort of knowledge that can influence the performance of a teacher: that used in thinking about the subject matter and the logical operations used in manipulating it. (p. 125)

The point is that a teacher is potentially better able to facilitate student achievement if he or she is aware of the type of knowledge--e.g., beliefs, concepts, and principles--being taught and the various logical operations--e.g., exemplifying, comparing and contrasting, describing, characterizing, classifying, inferring, explaining, justifying, abstracting, generalizing, and applying--used in manipulating each type of knowledge. With reference to understanding the various types of knowledge and the ways in which they can be manipulated, Smith (1969) observed:

> Because teachers do not now possess such understanding they frequently handle the subject matter of instruction in superficial ways. Consequently, class discussion often suffers from undue vagueness and ambiguity, from unfounded and unchallenged claims, from a failure to develop the significance of the content. (p. 126)

Hence, it seems that research related to (a) explicating various forms of mathematical knowledge, (b) explicating ways of presenting knowledge, and (c) identifying effective means of conveying different types of knowledge to teachers is essential for the development of a theory of teaching mathematics. Some groundwork for such research has already been completed by scholars who have theorized and explicated various types of knowledge, viz., mathematical concepts, principles, and skills. Some researchers have provided detailed and carefully defined descriptions of teachers' actions when teaching mathematical concepts, principles, and skills. Others have conducted empirical investigations dealing with the efficacy of various teaching strategies as defined in terms of the descriptive work of previous research.

The Necessity of Studying Teaching--Or-- Trying to Avoid Ostrichism

If various journals reporting research on teaching or on teacher education are reviewed, there are several observations one might make. Depending in part on how various research studies are categorized, it appears that relatively few studies have focused on teaching behavior, on the identification of principles for effective teaching, or on related problems in teacher education. A second striking factor is that studies investigating these problem areas have little common theoretical framework connecting them. This is not a criticism of the studies. In almost every case, the study is a solid piece of work and represents an extensive effort by the researcher. The lack of a common theoretical framework is not unique to research on teaching and is in fact characteristic of most research in mathematics education. (A concern for the lack of such a framework is essentially the raison d'etre of the Georgia Center for the Study of Learning and Teaching Mathematics.)

To some the paucity of research on teaching can be attributed to the belief that a teacher's effectiveness is primarily contingent upon attitudes and personality attributes of the teacher and influential factors in the child's home environment. Indeed, there is some empirical evidence to support these beliefs. Berliner (1975) noted that for subjects such as reading and social studies home influences are very powerful and could account for a substantial amount of the variance of student achievement. Berliner went on to point out that in subject areas not commonly learned at home, notably the sciences, socioeconomic conditions account for less variance than in reading, social studies, or language arts. Berliner (1975) concluded that there is more variance in acheivement of the sciences to be attributed to school and teacher effects (p. 16).

Still, some argue that teaching is an art and that efforts to subject it to scientific inquiry are futile. Davis (1967) subscribed to the following position:

> My proposition, obviously, is that the process of teaching is the practice of an art. It is not the application of a science in any presently meaningful sense of such a phrase: and the suggestion that it is must be labelled as, at best, a conspicuous instance of wishful thinking. (p. 38)

Gallagher (1970) made the following statements with respect to whether the artistry of teaching precludes scientific investigation.

> Is teaching an art? Indeed it is. Perhaps too much of one. Surgery was once too much an art and many people died as a result. Cooking is an art, and while few people die of it these days, drugstores do a thriving business in remedies for misbegotten creative culinary efforts. For when a set of skills is in a developmental stage where people say, "It is an art," they mean several things. First, that there are only a very few persons who have the skills that can identify them as highly effective practitioners, as "artists." Second, even these artists cannot give a systematic account of how they practice their art, and they are reduced to modeling their performance for those who would learn from them. But it is hard to imitate the true artist, and his genius too often dies with him.
>
> Those interested in the improvement of education and teaching would like to remove some of the mystery of the art of effective teaching through systematic study. (p. 30)

No thoughtful person would argue that mathematics teaching is not an art. The question is not whether mathematics teaching is an art but whether it is amenable to scientific inquiry. The position of the Teaching Strategies Project is that mathematics teaching can and will be improved through the elucidation of both analytic and empirical principles. Such principles can be identified through analyses of theoretical positions, descriptive investigations, and tightly controlled experimental studies.

Many artistic endeavors are subject to improvement as a result of scientific inquiry. Consider the field of athletics. Clearly an athletic event embodies the emotions of its participants. Indeed, emotions are sometimes a critical factor in an athlete's performance. Further, some athletes are more gifted, i.e., artistic, than others. Yet each athlete must adhere to some fundamental principles of performance, or the potential benefit of emotions and artistry will be greatly diminished. For the most part these principles are teachable and constitute items of instruction for coaches.

Probably no argument, however rational and eloquently stated, will change one's opinion on whether the teaching of mathematics should be studied in its own right. The controversy can and should be resolved only when evidence is available concerning the existence or nonexistence of principles for improving the teaching of mathematics. The conceptualization that began with Smith and Henderson has resulted in both evidence and promise for the continued search for such principles. Elaboration of that evidence and promise is one of the primary considerations of this monograph.

From one perspective there is little choice but to continue such research. That perspective relates to the advent and proliferation of competency based teacher education programs. Few would argue against the construct of competency based teacher education. (Is its complement incompetency based teacher education?) Yet the main hiatus in competency based teacher education is the dearth of reliable research that would identify maxims for teaching mathematics effectively. In view of the tens of thousands of students aspiring to become mathematics teachers and the profession's commitment to better teacher education programs, it seems reasonable and desirable to expand and coordinate research efforts. Brophy (1975) put it this way.

> Teacher educators and educational researchers need to pay more attention to the accumulation of a data base that would allow truly prescriptive teacher education to emerge. Propounding ideas on the basis of commitments rather than supportive data is unscientific to say the least, and blowing with the wind by propounding the latest educational fad is even worse. (p. 15)

Berliner (1975) has lamented the state of the art in competency based teacher education programs and teacher accountability systems. His comments below doubtless apply to some extent to all teacher education programs.

> Ostrichism is a common disease often afflicting education. Its etiology is a premature commitment to a particular educational movement. Behavioral symptoms include the practice of sticking one's head in the sand when problems appear, in the hope that the problems will go away. (p. 1)

The goal of the Georgia Center in general and of the Teaching Strategies Project in particular is to begin the arduous task of doing coordinated research. The justification for the workshop in teaching strategies, for which the papers in this monograph were commissioned, is the belief that the development of a theory of teaching mathematics would be facilitated by developing a coordinated research program. To date, most studies on teaching strategies have been dissertation studies. Although such studies can make significant contributions to knowledge about teaching mathematics, dissertation studies do not constitute the coordinated program of research that is necessary for the formulation of a theory. The Teaching Strategies Project intends to provide a catalyst for the identifications of a program of research.

A Perspective on the Papers

The papers in this monograph represent the thinking of individuals integrally involved in research on teaching mathematics. The purpose of this section is to mention and comment on particularly important points made by the various authors. Although the comments are, in a sense, specific to the Teaching Strategies Project, they also express concerns that apply to much educational research.

Smith's paper provides a context for understanding the historical development of the Project's current research. In addition to the interesting historical perspective, Smith summarizes several points that distinguish research in this project. One such point is the relevance of epistemology to the teaching act, that is, the nature of the knowledge being taught is a determining factor in how one should teach. This entails differentiating a teacher's interaction with content from his interaction with students. This distinction is a vital consideration for research.

Another point Smith makes is that since two variables--clarity and acceptance of pupil responses--have been identified as related to student achievement, there is promise that others can be identified. Perhaps so. "Clarity" is an interesting variable. The construct "clarity" has not been adequately characterized--its behavioral manifestations are not well defined and must be inferred as one observes teaching. The question arises as to how this variable is relevant to the Teaching Strategies Project. Is is possible, for example, to view clarity in the way that moves and strategies occur in teaching behavior? This topic will be explored in greater detail later in this paper.

Henderson's paper provides a basis for understanding and appreciating the evolution and contribution of pedagogical theory in teaching mathematics. His contribution, in concert with Smith's, identifies a unique characteristic of the project. The work of these two scholars and theorists provides the backdrop for the continuing research on moves and strategies in teaching mathematics.

Henderson makes several important suggestions concerning additional research. An especially intriguing suggestion concerns the diagnostic ability of mathematics teachers. As one reviews the literature on research on teaching, the large number and complexity of variables involved in the teaching process become increasingly apparent (in some ways painfully so). Probably no single dimension will account for effective teaching. If research on teaching points to anything it points to the reasonableness of this conjecture. This suggests that effective teaching may be related to a teacher's ability to recognize and react to specific classroom behaviors--in short, a teacher's diagnostic ability. The question then arises as to the extent to which a teacher's ability to identify and differentiate various interactive styles (teacher--student--teacher, etc.) can be described in terms of moves and strategies as these constructs are currently defined.

Henderson raised several questions in his oral presentation at the workshop that bear reiterating:

1. To what extent can research findings generalize across teachers, pedagogical approaches, or other subject areas?

2. To what extent should analytic work be done regarding the set of teachable objects?

3. Do research findings translate into textbook writings, classroom teaching practices, or both?

Since most research on teaching has been conducted by people not associated with mathematics education, the first question strikes at the heart of what we might claim to know about teaching mathematics. Henderson (in his oral presentation) stated that a theoretical approach to teaching should be based on the logic of various subject areas rather than on general psychological factors. This does not preclude the possibility, of course, that variables identified through educational research outside mathematics might have relevance to research on teaching mathematics. But it does raise the issue of what we can justifiably claim to know about the teaching of mathematics when our evidence is based on research not explicitly involved with mathematics.

The second question posed above focuses on the determination of the type of research that should be given priority in teaching strategies research. A great deal of analytic work has already been done in explicating the teaching of concepts, generalizations, and skills by Ginther (1965), Pavelka (1975), Semilla (1971), and Todd (1972). Other researchers have described how teachers justify knowledge (Wolfe, 1969) and how teachers help students organize knoweldge (Cooney & Henderson, 1972). For the most part these studies involved an interaction of logical considerations and analysis of teaching behavior.

At what point should such analytic work continue or yield to empirical investigations? Turner, in his oral presentation at the workshop, seemed to suggest that any additional analytic work on the models for teaching concepts, generalizations, and skills should be based on a need established by empirical evidence. Future research will probably not be entirely analytic or empirical. Rather these two types of research will likely emerge in concert with one another. The empirical research by Dossey and analytic work on explicating indicators of student learning represent two differenct yet mutually supportive research efforts.

The question of whether research findings translate into textbook writings, classroom teaching practices, or both is especially relevant to the research conducted by Dossey and Swank. Both investigators studied what Turner refers to as monadic strategies. The strategies in Dossey's treatments were expressed through programmed instruction and hence were carefully defined with greater confidence to the differences in the strategies employed. If a series of studies were to identify strategies that seemed particularly effective in a controlled setting, one might then ask whether they were also effective in the complex world of the classroom. Essentially the approach is one of doing "microcosmic" research in the hope that significant findings will emerge which will yield effective teaching principles in the "macrocosm" of classroom interactions. Dossey's investigation of the effects of various strategies across different types of concepts is particularly interesting.

Swank's primary objective, identifying effective teaching strategies, is the same as Dossey's. But the approach is different. Whereas Dossey examined specific and well-defined strategies, Swank examined the totality of the moves used in teaching concepts. In particular, he investigated the differential effects of a strategy having a "high" number of moves versus one having a "low" number of moves at two levels of classroom interaction. Should these "gross" strategies result in differential effects, then refinements of these strategies could be defined and investigated.

Swank did the teaching himself and used the school mathematics concept of function. These two features of the study made it more representative of classroom teaching. Yet Swank's strategies must still be classified as monadic since they were determined a priori and hence could not be completely sensitive to student responses. As Turner pointed out, a truly dyadic strategy complicates an experiment a great deal and can greatly escalate the cost.

The work of Dossey, Swank, and others doing similar studies can provide an empirical justification for pedagogical theory. Further, their research suggests several areas worthy of investigation. For example, Swank found that the treatment involving a high level of interaction facilitated achievement for higher ability students but not for lower ability students. Swank conjectured that the lower ability students may have been somewhat threatened by a higher amount of verbal interaction. This suggests that aptitude treatment interaction studies might be done in which the nature of the strategy is varied along with student characteristics.

Dossey suggested that the "power" of various moves be studied. Investigations might be conducted in a clinical setting where more detailed observations could be made regarding how various moves are received by students. Soviet educational psychologists use a methodology referred to as the "teaching experiment" in which the experimenter observes how students interact with the content as the teaching process proceeds. The "research product" is the qualitative aspects of student behavior observed by the experimenter. This methodology has a great deal of promise for the type of research Dossey suggested.

Turner's contribution suggests a way to structure research in teaching strategies. The need for such structure in all areas of educational research becomes increasingly more obvious as one reviews the literature. Merrill and Wood (1974) have also devised a scheme to structure research. Their model is based on four facets of instruction: learner aptitudes, subject matter content, instructional strategies, and instructional delivery systems. Although the two schemes are not entirely isomorphic, Merrill and Wood's facets relate closely to Turner's four primary domains: teachable objects, teacher actions, student attributes, and student indicators of learning.

Turner suggests not only a structure but also steps researchers can take to provide more precise experiments and coordination between experiments. He gives particular attention to problems related to defining and structuring elements within a domain and reducing known sources of variance within an experimental treatment.

One problem deserving attention is the construction of sampling frames for teachable objects. The large and heterogeneous class of teachable objects presents a real problem to anyone trying to generalize research findings. Without homogeneous classes of teachable objects, it is difficult to determine whether a finding involving a particular concept or principle is specific to that knowledge or whether it can be generalized to other concepts or principles. Analytic work involving the construction of homogeneous classes of teachable objects constitutes a desirable and in some sense, necessary project for those involved in research on teaching.

Another area worthy of consideration is the explication of different types of levels or indicators of student achievement for the various teachable objects. Turner's suggestion that indicators could be based on moves for teaching concepts, generalizations, and skills has in fact already been done--at least to some extent. Cooney, Davis, & Henderson (1975), in discussing the evaluation of student performance, use moves for teaching as the vehicle for assessing learning. The development of a hierarchy of moves and hence learning outcomes has not been done, however.

Turner makes the point that experimenters comparing treatments--for example, Dossey and Swank--should strive to make the various strategies equivalent in clarity and content. The less this equivalence is achieved, the more the comparison among treatments is compromised. For example, the determination of the differential effects of a characterization--exemplification--characterization (ECE) strategy versus an exemplification--characterization--exemplification (ECE) strategy is meaningful only if the strategies are equivalent in clarity and in the content of the moves comprising the C moves and the E moves. If one strategy is clearer than the other or if one group of C moves contains more information then another set of C moves, then the desired comparison is confounded. Concerns of this nature permeate much of educational research. Cronbach (1966) criticized experiments involving expository and discovery teaching because of biases in favor of the richness of the discovery treatments.

Turner also questions how one selects the strategies to be investigated. Obviously, the number of possible viable strategies is very large. In some studies , the strategies have been selected to correspond to the strategies teachers use in the classroom. This was the approach taken by Swank. Others--for example, Gaston and Kolb (1973)--selected strategies that typify various theoretical approaches to instruction. Still a rationale for the selection of moves and strategies is not yet altogether clear. Turner suggests that an alternative approach be considered: training teachers to use moves in their teaching. The comparison then becomes one of contrasting the effectiveness of teachers who have had such training with teachers who have not.

This approach has some appeal. For one thing, the ensuing treatments would then incorporate a dyadic definition of teaching. Perhaps a teacher who has been trained to use a variety of moves will be more effective than one who has not. Further, if certain student behaviors correlate with certain teacher behaviors, as Gregory and Osborne's (1975) work suggests, then one can ask whether a teacher's knowledge of moves is reflected in his or her students' behavior in class. If research follows this vein, and eventually it will, then the work of Retzer and others developing protocol materials will become increasingly significant.

Retzer's development and research activities can be characterized in terms of Turner's domains. For Retzer, the first domain (teachable objects) consists of knowledge of moves. The second domain (teacher actions) involves protocols and other types of delivery systems designed to teach the knowledge of moves. The third domain (student attributes) deals with the attributes of the trainee, e.g., whether the teacher is preservice or inservice. The fourth domain (student indicators of learning) deals with evidence that the trainee can in fact utilize knowledge of moves. The fifth domain (setting variables) constitutes the nature of the training program, the institution, and other aspects of the context of the delivery system.

Thus, research on delivery systems would constitute the selection of content (domain 1), a method of presenting that content (domain 2) with respect to the nature of the subjects in the sample (domain 3), and criterion measures for assessing outcomes (domain 4). In many respects the concerns raised above regarding research on identifying effective teaching strategies also apply to research on identifying effective delivery systems for knowledge of moves. Research of this type touches on an essential characteristic of competency based teacher education programs. Competency based programs are necessarily concerned with means of training prospective teachers to demonstrate certain desired and specified behaviors.

Research of the type Retzer suggests identifies the concern of how to determine if trainees have knowledge of moves and can utilize them in a teaching situation. To some extent one can use paper-and-pencil activities to ascertain knowledge of moves. Paper-and-pencil techniques, however, are inadequate to assess teaching ability. The problem is complicated partly because of what Retzer refers to as the "can do/will do" question. One can assess what a teacher can do under certain specified conditions. But whether a teacher will do in a classroom teaching situation what he can do is an open question. There is also the confounding issue of values.

A teacher might be able to use a certain technique but decide not to. This involves the teacher's value system. A teacher's values may be such that a certain move or procedure is consciously selected out. This presents a most difficult problem in evaluating teaching behavior. It is difficult to determine whether a teacher cannot use a move or strategy or whether he or she has elected not to use it.

Retzer points out that protocol materials are one means of teaching moves and pedagogical concepts. Gliessman emphasizes the necessity of developing protocol materials if research is to permit replication and interpretation. Hence, the development of protocol and training materials is and should be an important objective of this project. Because of the importance of developing protocols, several points raised by Retzer and Gliessman deserve mention. The first relates to the problems of producing materials, and the second relates to subsequent research involving the materials.

Gliessman emphasizes several key points to be considered in producing materials. One is to avoid materials which rely on multiple media. According to Gliessman, film materials that require a lot of printed material to explain their use are not likely to be used. Further, the materials should be brief and flexible. Extensive written materials generally result in vague and ambiguous concepts for the viewer. Of course, the pedagogical concept being depicted must itself be clearly defined with specific behavioral indicators or else the protocol is compromised from the start. Another point made by Gliessman is the desirability of producing materials with high technical quality and relatively free of noise. Materials that are well planned and conceptualized but are of poor technical quality generally will not be used by educators. Further, protocols must also be relatively free of noise when illustrating various concepts. For example, a technically superior film involving the use of counterexamples will not be well received if it also contains poor teaching practices. Viewers will likely focus on those poor practices to the exclusion of the concept being illustrated. An examination of various protocol films suggests that this is a nontrivial point.

One last point relates to the selection of a medium, e.g., film or film strips, for illustrating the concept. Gliessman urges that the full range of media be considered. In part, his advice is spurred by a cost-benefit question. Since high quality products can range anywhere from $400 for a ten-minute film strip to $12,000 for a ten-minute color motion picture, one must ask whether the benefits of a color motion picture warrant the extensive cost. The answer, in part, involves research on expected outcomes of using protocol materials. For example, in considering printed material such as transcripts or textbooks, audiotapes or movie films, one can ask which of these is superior in promoting observational skills, a teacher's use of moves, or other possible outcomes. At some point these questions should be dealt with by the producers of protocol materials.

Suydam's message, most welcome and worthy of consideration, is that research not specifically related to teaching strategies might reveal interesting questions and insights concerning our research. In fact, in the next section, additional issues in educational research will be raised that are relevant to the Teaching Strategies Project. In a sense, Suydam's comments cause one to reflect on the more global aspect of this project.

Suydam asks about the type of organization implicit in research on teaching strategies. To date most of the research on the differential effects of various strategies has been concerned with expository teaching. But the models do not dictate one type of teaching behavior. The models depict ways in which a teacher can "interact" with the content. Although the sequence of moves--assertion, instance, instance, instance, instance--suggests an expository strategy, the sequence--instance, instance, instance, instance, assertion--suggests a discovery approach or perhaps a laboratory approach where the instances are modeled in concrete objects. Hence, the moves can exemplify a variety of instructional modes.

The following sequence of questions raised by Suydam are basic to research on teaching strategies:

1. Does having teachers focus on specific types of language help students in achieving certain educational goals related to that language?

2. What language patterns do teachers use?

3. What is the effect of these patterns on students' performance?

The study that gave rise to these questions was conducted by Gregory (1972). Gregory sought to establish a relationship between a teacher's use of conditional logic and seventh-grade students' conditional reasoning ability. The report by Gregory and Osborne (1975) contained evidence that the relationship might exist. It is known that teachers vary in their use of moves for teaching concepts and principles. The pattern of these occurrences is still being investigated. If relationships between patterns of language and cognitive, or affective, outcomes can be identified, an educationally significant finding would be revealed. The pattern of the language might be ascertained by qualitative or quantitative aspects of moves and strategies. Or the pattern might be described by some other constructs, such as variability or clarity, that could be at least partially described in terms of moves or strategies.

Suydam raised other issues that also deserve consideration. Her emphasis on the role that questions play in teaching and in learning is well taken. Another issue is the need for researchers to communicate how terms are being used so better interpretations across studies can be made. Her suggestion that language patterns of teachers other than mathematics teachers be considered as a way of enlightening our insights into the teaching of mathematics is yet another issue worthy of consideration.

In the section "Last Reflections," Suydam observes that what is a good technique for some teachers is not always a good technique for others. As she suggests, humanistic concerns must always be an integral factor in viewing the totality of teaching. How various factors, affective and cognitive, weave together to articulate effective teaching principles, generic or specific, is a very complex and unsettled question. Amidst all of these complexities, one concludes using some empirical evidence, that the logical nature of a teacher's linguistic behavior is a potent variable for influencing learning.

Perspectives From Other Research

Earlier in this paper a case was made for the necessity and desirability of conducting research on teaching and on teacher education in general. In the preceding section, an attempt was made to provide a perspective on how research in the Teaching Strategies Project, as exemplified by the remaining papers in this monograph, can contribute to identifying effetive teaching principles. This section will discuss research on the teaching of mathematics and other subjects that has particular relevance to the work of this project.

It is fair to say that research has found out more about what variables are not related to effective teaching than about what variables are related to effective teaching. In reviewing the School Mathematics Study Group's (SMSG) research on teaching effectivensss, Fey (1969) noted that:

> Effectiveness of teaching using the SMSG materials is not significantly correlated with teacher's experience, collegiate courses and grades, or participation in professional activities. Most and least effective teachers were not differentiated by the amount of time they spent in preparation for teaching. There was only a weak indication that procedures in making assignments, explaining new material, conducting learning and thinking experiences relevant to previously assigned material, and evaluating and responding to student performance made a difference in the patterns of classroom behavior developed by effective and ineffective teachers. (p. 55)

What should be noted is that the variables identified above do not deal with teachers' classroom behavior, except for those measures involving Flanders' instrument. (It should be noted that other research involving this instrument has found certain patterns of teacher behavior to be associated with achievement.) Fey (1969), reviewing other studies in mathematics education, found that variables that describe what teachers are--as opposed to how they behave in the classroom--tended not to be related to achievement.

It seems reasonable to assume that research must focus on what the teacher does in the classroom. As mentioned above there has been very little of this kind of research in mathematics education. Exceptions, as indicated by the articles published in the first six years of The Journal for Research in Mathematics Education, consist primarily of studies on broad classifications of teaching, e.g., discovery versus expository or activity versus expository or studies in which a psycholgical construct such as learning hierarchies was utilized in instruction. One might have predicted the lack of research on mathematics teaching behavior. Holtan (1967) surveyed ongoing and proposed research activities in mathematics education. Of nearly 90 projects he identified, no more than ten and probably less than five could be classified as research on effective teaching behavior. Is there, then, some body of research relevant to the research in this project? There is, and some of this research is discussed below.

Promising Variables From Other Research on Teaching

Dunkin and Biddle (1974) and Rosenshine and Furst (1971) have provided impressive syntheses of research on teaching. Some of the variables in the studies they review appear to be potent--both empirically and intuitively--in predicting effective teaching behavior. Some of these variables involve affective and others cognitive aspects of teaching. In both reviews, two variables, viz., clarity and variability, are cited that might have particular promise for research on teaching mathematics.

However, these two variables are not easily defined. Dunkin and Biddle (1974) discuss what they call high inference variables and low inference variables. A high inference variable is one which is rather subjectively determined. That is, its behavioral manifestations are not well-defined. On the other hand, a low inference variable is one for which a behavorial definition is easier to obtain, e.g., a count of the number of questions asked. Usually high inference variables are more stable in teaching behavior and are of more interest to researchers. Clarity and variability are considered high inference variables.

Rosenshine and Furst (1971) cited seven investigations of the clarity of a teacher's presentation. Clarity was generally described in terms of whether the teacher's points were clear and easy to understand, whether the teacher had facility with the subject and could react to students in an intelligent way, and whether the cognitive level of the teacher's lesson was generally regarded to be appropriate for the students. In general, clarity accounted for a significant part of the variance of student achievement. Rosenshine and Furst (1971) noted that "In those studies for which simple correlations were available, the significant correlations ranged from .37 to .71" (p. 44). Rosenshine and Furst identified other studies in which the variables investigated were related to clarity and were significantly related to student achievement. Some of these variables were coherence of presentation, organization, and vagueness (negatively related to achievement). But it is difficult to ascertain what behaviors characterize clarity or the related variables even though ratings of these variables were relatively stable across occasions for a particular teacher and class.

According to Turner (1971), Bellack and Davitz investigated the congruence between the operation called for by a teacher's cue and the operation performed by the student. Turner noted:

> A distinctly bizarre outcome of the Bellack and Davitz study was that the type of rating reaction given by the teacher, classified as either positive or negative, remained constant at about 80 percent positive irrespective of the congruity or incongruity of the responses of the pupil. Thus, if the cue called for a definition and the pupil opined rather than defined, there appears to be a 4/5 probability that he would be positively reinforced if the teacher made a rating response at all. (p. 19)

Turner then asked how well teachers understand the logical dimensions of their cuing behavior.

Is clarity of presentation related to the congruity of the logical aspects of classroom interactions? Further, if this is a viable way to view clarity, can moves and strategies as presently explicated be used to help define clarity in terms of low inference variables? The teaching of mathematics seems to be suited to such an analysis.

Dunkin and Biddle (1974) reviewed several studies by Hiller and others in which the notion of vagueness of a teacher's presentation was considered. Vagueness was found to be negatively associated with pupil achievement. Further, vagueness seemed to be specific to teachers and not to lessons. This finding seems to support the findings related to clarity. Also, a teachers' knowledge of the subject tended to reduce the vagueness of a lesson. This finding seems to be inconsistent with the result of the research conducted with mathematics teachers using SMSG materials (Torrance & Parent, 1966).

Generally, a teacher's knowledge of a subject, as measured by an achievement test, has not been found to be related to student achievement. Perhaps, however, there is a more fruitful way of defining knowledge of a subject. Turner (1971) made the following suggestion:

> Perhaps the relevant evidence for whether a teacher "knows the subject" does not lie in whether he can correctly answer the items on a test so much as it lies in his reactions to pupil responses to the cues he himself has emitted--in short, in the kind of performance standards he employs with respect to substantive responses from pupils. (p. 20)

This leads to the consideration, then, that the logical dimension of the teaching act as depicted by moves developed by Henderson and others might serve as a means of defining a teacher's knowledge. Consider a mathematical concept. Can a teacher identify examples or nonexamples of the concept? Can necessary or sufficient conditions for objects to exemplify a concept be identified? Can the teacher present a counterexample given a false generalization by a student? Such questions could begin to provide a basis for defining a teacher's knowledge of content.

A second promising variable is that of variability of a teacher's teaching behavior. Rosenshine and Furst (1971), in reviewing research on teaching, considered this variable to be quite promising. As with clarity however, variability is a high inference variable for which low inference behaviors indicating variability have yet to be identified. Indeed, variability has been defined in many ways, ranging from the variety of materials used to cognitive variability in classroom discourse. Dunkin and Biddle (1974), in reviewing studies on logical variability, noted that Furst found that teachers' logical variability--that is, using analytic and evaluative comments in contrast to largely empirical statements--was positively related to pupil achievement.

Fey (1969), in reviewing the SMSG studies on teacher effectiveness, pointed out that the productive thinking ability of teachers was a significant predictor of teaching effectiveness. Fey noted:

> The most effective teachers produced more ideas about indications of success or failure in their teaching, causes of success or failure, and alternative ways of teaching course concepts. (p. 55)

It is not clear how far the notion of productive thinking as described above is consistent with the notion of variability. What seems intuitively clear, however, is that a teacher who can be flexible and insightful concerning factors that contribute to the effectiveness of a lesson is better equipped to promote learning than one who is not. Substantial work needs to be done to investigate relationships between particular classroom situations and how teachers can react cognitively (as well as affectively, of course) to those situations. Again the question arises as to whether knowledge of moves and strategies can increase a teacher's variability in teaching.

The variables of clarity and variability were identified by Rosenshine and Furst (1971) as two of the five most promising variables based on empirical evidence for predicting student achievement. (The remaining three were enthusiasm, task oriented or business like behaviors, and student opportunity to learn.) The two variables were chosen for review here, in part, because empirical evidence suggests they are related to achievement and, in part, because of their appeal to related research in teaching strategies. Henderson, in this monograph, points out the desirability of investigating variables within the domain of specific content areas. The analytic nature of mathematics may permit research on clarity and variability with a precision not possible in other subject areas.

Several questions come to mind as a result of the above discussion. They have been suggested previously but are reiterated here for emphasis.

1. If one is willing to accept that clarity and variability have potential for predicting success in mathematics achievement, how can this relationship best be investigated and established?

2. To what extent can moves and strategies be used to define the cognitive variables of clarity and variability?

3. Can a teacher be trained to exhibit a desired level of clarity and, if the answer is yes, how can this training be carried out?

Conclusion

In this paper, issues on research on teaching have been raised and considered from various perspectives. The goal of research concerning these issues is the identification of effective teaching principles. There are various approaches that can be used to identify such principles. Studies which are carefully controlled and focus on a monadic definition, such as Dossey's, exemplify one such approach. Another is the use of descriptive studies which explicate various aspects of the teaching act. Medley (1973) emphasizes the need for researchers to capture the dynamic aspects of teaching. A teacher's competence is related more to when a teacher uses a particular technique than the mere fact that it occurs. To illustrate his point, Medley (1973) made the following analogy:

> Medical research does not concern itself with whether the best doctors use penicillin more often than, say, cortisone; it concerns itself with what penicillin is good for, and what parameters or conditions determine its effects, as well as with what cortisone is good for and what parameters determine its effects. (p. 44)

Descriptive research and tightly controlled research do not have to be disjoint. Descriptive work serves the role of identifying variables and relationships worthy of investigation. Consider, for example, the teaching experiments of the Soviet researchers. Menchinskaya (1969) investigated the effects of varying irrelevant and relevant attributes in teaching mathematical concepts. Zykova (1969) observed the effect of various visual representations of examples and how they affect acquisition of mathematical concepts. These studies are primarily descriptive in nature, focusing on highly qualitative aspects of teaching and learning. Findings of such research can influence the work of highly controlled studies, such as those of Shumway (1974) or Dossey and Henderson (1974), that are relevant to research on teaching strategies.

There is also a need to determine if and how teachers can be trained to utilize moves or strategies or to demonstrate some desired behavior such as clarity or variability. Presently, pedagogical theory on teaching mathematics does not consist of empirical generalizations that can be asserted with great confidence. This situation is not unique to knowledge about teaching mathematics. It does raise, however, a question concerning the basis on which teacher education programs can be constructed. Given the present state of the art, mathematics teacher education programs must be content with closing the gap between what is believed desirable and what is actually the case.

The growing national movement to hold individuals and institutions accountable for their actions makes it more desirable than ever to investigate principles of effective teaching methods. Teaching is designed to promote learning. But learning can be caused or inhibited by factors other than the actions of a teacher. Implicit in Retzer's paper in this monograph is the belief that teachers should take responsibility for defining those actions for which they are willing to be accountable. As an analogy, Retzer (1976) noted that in the case of physician's accountability, one speaks of malpractice not malhealing. Moves and strategies are one area worthy of consideration in trying to define teacher accountability.

One other observation should be made. It relates to the concern, raised in part by Suydam, that teaching is a very complex phenomenon involving a large number of variables. In short, teaching can not likely be profitably conceived as a unidimensional behavior. Fey (1969), whom Suydam cites, says it quite aptly:

> The question of predicting teacher effectiveness is not simply answered by direct measurement of obvious variables, but must be viewed as a complex interaction of several interrelated classes of variables. (p. 60)

Clearly, variation in sequences of moves as presently conceived can not by themselves singularly and totally account for teacher effectiveness. Two teachers using exactly the same strategies are likely to get different results if affective variables differ greatly. Yet the question remains as to whether, on the whole, the logical nature of moves is a significant variable in viewing instruction. Surely all else being equal rapport is an important ingredient for effective teaching. But it is highly doubtful that rapport alone--or any affective variable--can account for achievement where extensive differences between teachers exist on cognitive measures. Thus, the question remains and deserves consideration: "To what extent can moves and strategies account for variance in effective teaching?"

Finally, the reader may have noticed that this introduction has raised far more questions than it has answered. Both the introduction and the remainder of the monograph are intended to stimulate and challenge those people interested in research on teaching mathematics. If the intended stimulation and challenge leads to research on teaching strategies, then the unanswered questions will have served their purpose.

References

Berliner, D. C. Impediments to the study of teacher effectiveness (Tech. Rep. 75-11-3). San Francisco: Far West Laboratory, 1975.

Brophy, J. E. Reflections on research in elementary schools. Paper presented at the conference on Research on teacher effects: An examination by decision-makers and researchers, University of Texas, Austin, Texas, November 1975.

Cooney, T. J., Davis, E. J., & Henderson, K. B. Dynamics of teaching secondary school mathematics. Boston: Houghton Mifflin, 1975.

Cooney, T. J., & Henderson, K. B. Ways mathematics teachers help students organize knowledge. Journal for Research in Mathematics Education, 1972, 3, 21-31.

Cronbach, L. J. The logic of experiments on discovery. In L. S. Shulman & E. R. Keislar (Eds.), Learning by discovery: A critical appraisal. Chicago: Rand McNally, 1966.

Davis, R. B. Mathematics teaching--with special reference to epistemological problems. Journal for Research in Education Monograph (No. I) 1967.

Dossey, J. A., & Henderson, K. B. The relative effectiveness of four strategies for teaching disjunctive concepts in mathematics. Journal for Research in Mathematics Education, 1974, 5, 6-19.

Dunkin, M. J., & Biddle, B. J. The study of teaching. New York: Holt, Rinehart & Winston, 1974.

Fey, J. T. Classroom teaching of mathematics. In J. W. Wilson & L. R. Carry (Eds.), Reviews of recent research in mathematics. Studies in mathematics (Vol. 19). Stanford, California: School Mathematics Study Group, 1969.

Gallagher, J. J. Three studies of the classroom. In J. J. Gallagher, G. A. Nuthall, & B. Rosenshine (Eds.), Classroom Observation. American Educational Research Association Monograph Series on Curriculum Evaluation, Monograph No. 6. Chicago: Rand McNally, 1970.

Gaston, J. A., & Kolb, J. R. A comparison of three strategies for teaching a selected mathematical concept to students in college algebra. *Journal for Research in Mathematics Education*, 1973, *4*, 177-186.

Ginther, J. L. A study of definitions in high school mathematics textbooks (Doctoral dissertation, University of Illinois, 1964). *Dissertation Abstracts*, 1965, *25*, 4574. (University Microfilms No. 65-00818)

Gregory, J. W. The impact of the verbal environment in mathematics classrooms on seventh grade students' logical abilities (Doctoral dissertation, The Ohio State University, 1972). *Dissertation Abstracts International*, 1972, *33A*, 1585. (University Microfilms No. 72-27, 014)

Gregory, J. W., & Osborne, A. R. Logical reasoning ability and teacher verbal behavior within the mathematics classroom. *Journal for Research in Mathematics Education*, 1975, *6*, 26-36.

Henderson, K. B. A theoretical model for teaching. In R. T. Hyman (Ed.), *Contemporary thought on teaching*. Englewood Cliffs: Prentice Hall, 1971.

Henderson, K. B. *The use of pedagogical models*. Paper presented to the Special Interest Group in Mathematics Education at a meeting of the American Education Research Association, Chicago, April 1972.

Holton, B. Some ongoing research and suggested research problems in mathematics education. In J. M. Scandura (Ed.), *Research in Mathematics Education*. Washington, D. C.: National Council of Teachers of Mathematics, 1967.

Hyman, R. T. Teaching: Triadic and dynamic. In R. T. Hyman (Ed.), *Contemporary thought on teaching*. Englewood Cliffs: Prentice Hall, 1971.

Medley, D. M. Closing the gap between research in teacher effectiveness and the teacher education curriculum. *Journal of Research and Development in Education*, 1973, *7* (1), 39-46.

Menchinskaya, N. A. The psychology of mastering concepts: Fundamental problems and methods of research. In J. Kilpatrick & I. Wirszup (Eds.), *Soviet studies in the psychology of learning and teaching mathematics* (Vol. 1) Stanford, California: School Mathematics Study Group, 1969.

Merrill, M. D., & Wood, N. D. *Instructional strategies: A preliminary taxonomy*. Columbus, Ohio: ERIC information Analysis Center for Science, Mathematics, and Environmental Education, 1974.

Pavelka, E. B. Moves, strategies, and models of a concept venture in secondary school mathematics (Doctoral dissertation, The University of Illinois, 1974). Dissertation Abstracts International, 1975, 35A, 7183. (University Microfilms No. 75-11, 667)

Retzer, K. A. Personal communication, January 1976.

Rosenshine, B., & Furst, N. Research in teacher performance criteria. In B. O. Smith (Ed.), Research in teacher education: A symposium. Englewood Cliffs, N. J.: Prentice Hall, 1971.

Semilla, L. Z. Moves and strategies in a principle venture in secondary school mathematics (Doctoral dissertation, University of Illinois, 1971). Dissertation Abstracts International, 1972, 32, 4495. (University Microfilms No. 72-07054)

Shumway, R. T. Negative instances in mathematical concept acquisition: Transfer effects between the concept of commutativity and associativity. Journal for Research in Mathematics Education, 1974, 5, 197-211.

Smith, B. O. Teachers for the real world. Washington, D.C.: The American Association of Colleges for Teacher Education, 1969.

Todd, H. W. Moves and strategies in a skill venture in secondary school mathematics (Doctoral dissertation, University of Illinois, 1972). Dissertation Abstracts International, 1973, 34, 654. (University Microfilms No. 73-17449)

Torrance, E. P., & Parent, E. Characteristics of mathematics teachers that affect student learning (Cooperative Research Project No. 1020). Minneapolis: Minnesota School Mathematics and Science Center, Institute of Technology, University of Minnesota, 1966.

Turner, R. L. Conceptual foundations of research in teacher education. In B. O. Smith (Ed.), Research in teacher education: A symposium. Englewood Cliffs, N. J.: Prentice Hall, 1971.

Wolfe, R. E. Strategies of justification used in the classroom by teachers of secondary school mathematics (Doctoral dissertation, University of Illinois, 1969). Dissertation Abstracts International, 1969, 30A 1064-1065. (University Microfilms No. 69-15, 422)

Zykova, V. I. Operating with concepts when solving geometry problems. In J. Kilpatrick & I. Wirszup (Eds.), Soviet studies in the psychology of learning and teaching mathematics (Vol. 1). Stanford, California: School Mathematics Study Group, 1969.

Teaching Strategies: Historical and Contemporary Perspectives

B. Othanel Smith

University of South Florida

The term "strategy" comes from the Greek word "strategia" meaning generalship. It consists of a combination of two words, "stratos" meaning army and "agein" meaning to lead. It pertains to the art of maneuvering troops or ships, and it is used to denote the science of military command in conflict with an enemy. In a more general sense, the word "strategy" is used to denote a particular plan or method of reaching a goal. Such a plan or method is an instance of strategy and might appropriately be called a strategem, although we shall not use the singular form.

Because of its military origin, we think of strategy in the context of conflict: the strategy of international politics, of a political campaign, of a military campaign, of labor in its struggle with management. The behavior involved in planning and executing a strategy, such as the strategy of a military campaign, must presuppose the position of the enemy, his resources, his likely moves, and so on. The best strategy depends upon what the enemy will do as well as his advantages and disadvantages.

Pedagogical Uses

Some persons decry the use of the term "strategy" in pedagogy. It is easy to believe that naive historians a generation hence will write of the military mentality of those of us who use such terms.

Be that as it may, there are some useful parallels. The classroom involves a teacher and a number of pupils in situations where objectives are to be attained through the actions of the teacher, often in the face of such resistances as conflicting motivations and cognitive strains. How are such situations similar to the ones faced by a general? They both have ends to be attained and means are used to attain them. They are alike, also, because the teacher and the pupils do not always see alike, and hence the situation often exhibits differences of motivation, information, and opinion. Then, too, pupils must be organized and skillfully managed if goals are to be realized. Also, like the military situation, the parties are in no sense peers. The authority relationship in the teaching situation is one in which the teacher is always in a position of greater influence than the pupil. But the similarities end here. The teaching situation is different in at least two aspects. For one thing, the outcomes are beneficial to the pupil, and they are no loss to the teacher. For another thing, both teacher and pupil typically become engaged in cooperative activities that take the pupil into the domain of knowledge possessed by the teacher.

Pedagogically, strategy refers to a set of actions that serves to attain certain results and to guard against others. From a general standpoint, strategies are used to induce pupils to engage in verbal exchange, to insure that certain points will be clarified, to reduce the irrelevant or wrong responses, to cooperate in activities, to assure that objectives will be attained, and so on.

Apparently "strategy" is a label for what has historically been called method or procedure. The gain from changing terminology is not so apparent. Changing the name of a horse does not enable him to win more races. The question of whether we are talking about the same horse or a different one is not wholly clear. The term "strategy" is being used so fast and loose by so many persons that one hardly knows what is being referred to. We hear such expressions as the "strategy of measurement," "curriculum strategy," "control strategy," and we are told that sequencing and pacing of learning are strategies. All that aside, why abandon the word "method?" There may be no good reason. However, "method" is closely identified with a priori approaches to the improvement of teaching. When we began to look at teaching behavior as worthy of study in its own right, to analyze it, and to seek ways of improving it as such, we needed new words to designate what we found. So we began to use such labels as strategy, moves, or tactics. The concepts designated by these labels are similar to the old concepts of method and techniques, but they are also different. The differences and similarities can be readily seen in a historical context.

Grand Strategies of the Recent Past

Up until about three decades ago, the improvement of instruction consisted largely of attempts to impose upon teaching behavior patterns derived from psychological and philosophical concepts. In this century, there are four notable examples of this approach, all rooted in the Darwinian mentality. The Darwinian view is that the world is a situation in which living things either adapt themselves or they perish. This adaptation takes place on a grand scale in which shifts in the environment bring about the necessity for new ways of existing. These new ways are determined by variation, natural selection, and a resulting adaptation of the species to the new situation.

In one way or another the patterns of teaching behavior derived from psychology and philosophy reflect this cosmic model of adaptation and survival. The maze and the puzzle box contrived by Thorndike to study the learning of animals were miniatures of this cosmic pattern. Deprived of food in the puzzle box the cat's survival depended upon his escape. He did so by trial and error. The successful trials were retained, and the errors were stamped out. The process of learning was that of natural selection among the variations of behavior, resulting in adaptation and survival.

To make sure, Thorndike (1965) ruled out the possibility of imitation by placing the cat in a position to see other cats escaping from the same puzzling situation. He ruled out instruction by putting the cat through the movements which were necessary for escape. In neither case did the cat learn to escape from the box.

Thorndike (1909) thus concluded that the cat learned to extricate itself neither by intelligence nor by instruction. This conclusion reflected his general position that

> in life and in mind the same cause will always produce the same effect, that the bodies and mind of men are a part of nature, that their history is as natural as the history of the stars, their behavior as natural as the behavior of an atom of hydrogen.

To improve teaching, according to this view, is to discover the natural conditions of learning and to shape the teacher's behavior to these conditions. Thorndike (1914) stated the natural conditions of learning in his three laws: the law of readiness, the law of exercise, and the law of effect. The law of readiness entails some conditions of imbalance between the organism and its environment, for example, hunger, as in the case of the cat in the puzzle box. The law of exercise required that the learner either practice what was planned for him or, as in the puzzle box, engage in random activities. The law of effect entails dissatisfaction from failure or satisfaction from successful trials.

This pattern was first imposed upon the teaching of reading. If the teacher provided the child with chances to read while in a state of readiness, the child would make attempts to read and by the effect of rewards would be encouraged to continue his efforts and thereby learn. Even today the concept of readiness is basic to the teaching of reading. A number of attempts were made to apply Thorndike's laws of learning to teaching in general. By and large his conception of how learning occurs has permeated our thinking about teaching to such an extent that we are often unaware that we are persuaded by his views.

A second notable example of the imposition of a pattern upon teaching behavior is to be found in the work of Skinner (1968). Trial and error left little opportunity for the experimenter deliberately to shape the response which the cat makes to the puzzle situation. It is to Skinner that we must attribute the insight that the cat's behavior in the situation could be influenced by the experimenter. One of the ways in which this could be done was to reward those behaviors which were in the direction of success and not to reward those that were inappropriate. The experimenter could thereby shape the behavior of the animal by successful approximations toward the desired pattern. Skinner thus refined Thorndike's laws of exercise and effect. He substituted for these laws the concept of shaping the behavior of the learner by successive increments in the direction of the ultimate behavior.

This pattern of teaching was tried out through programmed instruction which involved incremental learning by the influence of planned reinforcement. Today the primary emphasis is upon what is frequently referred to as behavior modification where teaching becomes a matter of reinforcing desired behavior and not reinforcing that which is undesired. Pupil behavior is thus at last framed in accordance with the natural conditions of learning, and the successful teacher is deemed to be one whose behavior institutes classroom circumstances in conformity with these conditions.

Two other notable examples of attempts to impose a pattern upon teaching behavior are from the field of philosophy. These two patterns came also from the Darwinian orientation. The first is found in the works of Dewey (1910) about the beginning of this century. Dewey had come to the view that man's intellectual activities were a response to some sort of perplexing situation. Man coped with a situation by analyzing it and formulating a plan for dealing with it. Dewey called this plan hypothesis. After its formulation the plan would be tried out, first in imaginative rehearsal and after that by action. If the consequences of the plan turned out to be a satisfactory resolution of the situation, the problem was deemed to have been solved.

The parallel between Thorndike and Dewey's thought is fairly evident. Instead of the situation inducing a state of readiness, as in the case of hunger, Dewey substituted the notion of a perplexing situation in which a felt need emerges. For the notion of trial and error, Dewey substituted the idea of a hypothesis or plan of action tested in preliminary fashion by mental rehearsal of consequences rather than by overt trials. If the hypothesis survived the rehearsal, it could then be tested in action. Like Thorndike, Dewey's general notion was that of continuity in nature. He held that lower forms of life, even one-celled forms of existence, reflected in their behavior the pattern of inquiry in man himself.

Witness here again the view of man as adaptable. But in Dewey's view, man adapts himself through the use of intelligence rather than through sheer trial and error.

In his Democracy and Education Dewey (1916) elaborated his theory of how man thinks into a pedagogical method often referred to as the inquiry method of teaching. It remained for Kilpatrick (1925), a few years later, to formulate Dewey's notions into a generalized method which he called the broader method, more popularly known as the project method. This method of teaching consisted of four elements: purposing, planning, executing, evaluating. In Kilpatrick's view, pupils are in a situation that induces purposing. Once they have decided upon their purposes, they plan, that is, they decide how these purposes can be realized. When their plans have been formulated, they then carry them out. At the end of their work they turn around and look at what they have done and evaluate it.

It is easy to see from this general pattern of pupil activity that the work of the teacher consists more in helping pupils purpose, plan, execute, and evaluate their work than in imparting information to them. To be sure, the teacher is to be helpful and to supply information at points where pupils ask for it, or where the teacher thinks it desirable. But giving information is to be handled stringently lest the pupils lose their momentum as autonomous learners.

It is important to note that the idea of a situational base for learning, found in Thorndike, Dewey, and Kilpatrick, and only to a lesser extent in Skinner's work, has not only permeated our common sense but has also led to a theory of learning outcomes (Kilpatrick, 1925). In a situation there are many stimuli. Among them are primary stimuli which induce behavior resulting in outcomes identified with the knowledge being studied. There are many other stimuli in the situation that evoke behavior unrelated to the main or primary outcomes of study. This behavior leads to attitudes or affective learnings.

Kilpatrick (1925) made a great deal of this point. It became the basis of his concept of concomitant learnings. These learnings are those which occur along with primary learnings or the cognitive outcomes associated with the discipline under study. For example, at the time that a pupil is acquiring the learnings of history or science, he is also learning to like or to dislike the teacher, to like or dislike the subject, to like or dislike school, and so on. These concomitant learnings have been rediscovered and referred to nowadays as unplanned outcomes. They are central features of Kilpatrick's conception of the broader method of teaching set forth fifty years ago.

A second notable example is the Morrisonian (1926) theory of teaching. Morrison, too, worked in the shadow of Darwin. To him learning was also a way of adjusting to the world. The knowledge that man has accumulated and the moral notions he uses to guide his conduct are forms of adaptation which have enabled man to survive in the world as it is. As Morrison (1934) put it;

> So long as education is necessarily a matter of adjustment by process of adaptation and so long as the object of adjustment is a world which is common to all mankind, it follows that it must be possible to construct a curriculum which is objectively valid. (p. 49)

The objectives which constitute the curriculum are always either attitudes or acquired abilities. He divides attitudes into two categories: attitudes of understanding and attitudes of appreciation. Attitudes of understanding are to be found primarily in the sciences and attitudes of appreciation primarily in the arts. Abilities are found in the uses of language, in bodily activities, and so on. To attain these objectives is to acquire a personality adaptation. Once a personality adaptation has occurred it is never lost. In Morrison's view when the learner has mastered an ability or an attitude of understanding or an attitude of appreciation, he never forgets or loses it. This notion is fundamental to his mastery formula for teaching.

It is interesting to note that Morrison distinguishes between performance and learning outcomes and experience and learning outcomes. His view is that one does not learn an experience but that he learns from experience. In short, learning is not behavior but an adaptation of personality. This adaptation controls subsequent behavior. He simply took the Darwinian notion of adaptation to an environment as the basis of survival and translated it into an adaptation of personality as the outcome of a learning cycle.

The pattern of teaching behavior corresponding to the cycle of learning is what he refers to as systematics teaching or his mastery formula. The mastery formula is: pretest, teach, test, diagnose, and teach again. Repeat to the point of mastery.

These four conceptions have dominated attempts to improve teaching throughout this century. From time to time they have faded away only to reappear under new names. The theory of the project method is enjoying a recrudescence under the name of the open classroom, and the Morrisonian theory is coming to the front again under the stimulation of the concept of mastery put forth by Bloom.

These theories are grand strategies if strategy is defined as a plan for coping with any affair. Each in its own way is a formula for coping with the problems of teaching regardless of both content and pupils. The appeal of these strategies is almost overwhelming. They embrace a considerable measure of truth and their effectiveness with certain pupils can hardly be doubted, but as universal strategies they are not easily defended. Yet, they intoxicate the pedagogical mind, create doctrinal disputes, and lead repeatedly to schisms in the teaching profession.

Empirically Based Strategies

About 25 years ago, a few researchers became disenchanted with the grand strategies of the masters. They began to look at teaching as a form of behavior to be studied in its own right.

Teaching, like all the professions, has its origin in the practices of primitive man. It is found in man's first efforts to pass on his knowledge and abilities to the young just as the work of the physician is foreshadowed in the primitive medicine man's care of the afflicted. Primitive crafts that ultimately become professions develop through the accretion of knowledge and skill. To understand a craft is to possess the prime conditions for its development. Exploration of the phenomenon of teaching is now in its early stages. This late start is attributable partly to the fact that all along we assumed that if we understood the phenomenon of learning, we would thereby know how to teach; and partly to the fact that we were taken captive by the belief that teaching could be derived from the philosophic and psychological disciplines. As a result, the study of teaching in its own right received but little attention prior to the 1950's.

About this time teaching began to be perceived as a form of social behavior having its own forms, irregularities, and problems. In addition, teaching takes place under relatively constant conditions, namely, time limits, authority figure, student ability, institutional structures, and social expectations. Given these perceptions, naturally the development of a rational plan for the improvement of teaching would begin with an understanding of teaching behavior itself.

It is also evident that to apply any theory one must understand the phenomenon to which it is to be applied. It is just as necessary for one to understand what teaching is as it is for him to understand the concepts and principles which he applies to improve it. The turn about in efforts to improve teaching behavior is due in part to the belief that one must first understand the dimensions of such behavior before he can think realistically about its control through the application of principles.

During the last three decades, scores of studies have been made in an attempt to determine the dimensions of teaching and in a few cases to relate these dimensions to pupil achievement. These studies fall roughly into two categories: those which use a checklist or rating scale and those which rely upon recording instruments for data.

There are two notable cases among those who used inventories and also related dimensions of teaching to achievement of pupils. Ryans (1960) in his study correlated categories of teaching behavior to pupil achievement. He found that those teachers who were systematic and businesslike were more successful in inducing achievement than teachers who were unsystematic and rambling. The second example is the Flander's (1970) study. This research shows that teachers who make use of pupil ideas and who accept the pupil's feelings and encourage them have higher achievement among their pupils than teachers who are more directive.

These studies tell us the good and bad practices that occur normally in classroom situations. This information enables us to emphasize the effective practices and to eliminate those less successful. These studies thus enable us to refine conventional teaching behavior by eliminating ineffectual aspects of it. They do not impose upon teaching behavior a new pattern, but instead these studies maintain the old pattern and refine it to make it more conducive to pupil achievement.

There have been a number of attempts, some of them successful, to show the relationship between the dimensions of teaching, determined by analysis of tape recordings, and the achievement of pupils. However, the emphasis has been upon describing teaching behavior as such rather than to define it into more effective patterns of behavior.

Both approaches to the study of teaching have revealed a number of variables related in more or less degree to student achievement. Among these variables are clarity, variability of activities and levels of learning, task-oriented or business-like behavior, uses of student ideas and general indirectness of teacher behavior, uses of structuring comments and types of questions (Rosenshine & Furst, 1971). These are modest achievements, but they indicate that this approach to the improvement of instruction is viable and give ample justification for continuing it with hope and enthusiasm.

Some of these studies (e.g., Flanders, 1970; Ryans, 1960) retain one feature of the grand strategies. They assume, as do the grand strategies, that all of the different kinds of learnings are to be taught by the same strategy; that is to say, strategies which apply to the teaching of concepts apply also to the teaching of rules, cause-effect relationship, appreciations, morals, and psychomotor skills as well as to all disciplines.

My own studies have led me to the view that teaching behavior varies with the discipline and the nature of the learning product. However, like Thorndike and others, we began with the notion that patterns of behavior could be imposed upon teaching. This is manifested in our early belief that certain logical operations could be imposed upon teaching behavior; that if this were done the critical thinking of pupils would be enhanced. Professor Henderson (1953) and I, being of the same persuasion, collaborated with some high school teachers in a project to determine whether or not critical thinking could be improved by teaching the logic of the disciplines along with their content.

As our work progressed it became evident that the logical operations we were asking teachers to perform did not fit into their patterns of behavior. Yet from cursory classroom observations, it seemed to be evident that some logical operations were in the teachers' performance. They did define terms, explain events, and describe phenomena.

This apparent congruity between teacher performance and the operations we were attempting to impose upon teaching led me to undertake a sequel to the critical-thinking study. Our first task was to tape record the discourse of teachers and pupils and to analyze it into categories of logic.

We found that teachers did in fact perform a number of logical operations with their pupils. They defined, designated, classified, compared and contrasted, explained, evaluated, opined and inferred conditionally (Smith & Meux, 1970). They did these very loosely and imperfectly, typically engaging in only a part of an operation. But even more important, these operations were not performed as ends in themselves but as elements in a larger pattern of behavior. This fact seemed to account partly for their incompleteness and the resistance of teachers to the performance of logical operations as they taught their disciplines.

We then conducted a further analysis of the classroom discourse to identify the maneuvers within which logical operations are embedded (Smith, Meux, Coombs, Nuthall, & Precians, 1967). After trying a number of approaches, we finally broke the discourse into units called ventures. We used as a criterion of a venture that the discourse leads to a meaningful object, that it have import. We thus classified the ventures in accordance with the meaningful objects embedded in them. This procedure resulted in a number of different types of ventures: concept ventures, appreciation ventures, cause-effect ventures, reason ventures, value ventures, rules ventures, procedural ventures, and particular or factual ventures. These varied in length from a few lines to several pages. Professor Henderson has referred to the objects embedded in these ventures as "teachable objects," and we shall use that term here.

In the course of teaching these objects, teachers and pupils verbally manipulate the content of instruction. We called these "manipulations moves." Consider an example from a causal venture. There are three classes of moves: those which describe causes, those that describe effects, and those which relate the two. An example that relates cause to effect is taken from a social studies class. It is a chaining move. Such a move links a cause to an effect by describing conditions leading to the effect. The following move links the discovery of gold in Alaska to increased prosperity in the United States:

Pupil: The value of the gold means that the prices of corn and wheat go up.

Teacher: All right. The fall in the values there affected prices so the farmer profits.

The fall in the value of gold is noted as a consequence of the discovery of gold (the cause). This consequence, in turn, is noted as causing an increase in the price of farm products. This price increase is part of what is meant by prosperity. In this move there is a complete chain, though poorly described.

We identified the moves involved in each type of venture. The character of the moves varies with the kind of venture. The moves in a causal venture are quite different from those in a concept venture or in an appreciation venture. When a number of ventures of the same type are analyzed across disciplines, it turns out that each type has more than one pattern of moves depending upon the discipline. These patterns constitute the large maneuvers which we refer to as strategy, and the moves they contain represent elements of logical operations. In our view, therefore, one does not speak of a strategy of teaching in general but of strategies appropriate to each teachable object in each discipline.

Empirical studies of teaching have led us to conclude that there are two types of strategies. One type is content bound. The strategy for teaching concepts, for example, is not the same as the strategy for teaching cause-effect relationships or rules or for that matter any other of the teachable objects. The strategy varies also with the discipline. The second type of strategy is content free. Flanders' (1970) acceptance behavior is an example of this type. Put in general terms, it is a strategy of ingratiating and involving the student by accepting his feelings and using his responses in a constructive and acceptable manner. This kind of strategy, being content free, is compatible with the grand strategies of the masters. General strategies, such as that of Flanders, are for interacting with pupils. The content bound strategies are for interacting with the content of instruction. Both strategies are essential in the teaching process. We hypothesize that an improvement in the performance of either one will influence the outcome of classroom activities.

Summary

What have we learned from the last 25 years of research on teaching, and what difficult matters have not been touched upon? We summarize these as follows:

1. We are beginning to see how the performance of the teacher is dependent upon the form of knowledge and the discipline he teaches. Some authorities have insisted from time to time that methods of instruction should be related to the content of instruction. This relationship was explored briefly by Thorndike and others during the first quarter of the century, but only now are we beginning to see that laws, concepts, and procedures, for example, are taught in quite different ways. Recent research has also emphasized the cognitive difference among disciplines. For example, empirical sciences contain laws, concepts, and procedures as forms of knowledge. Mathematics contains no laws; history has law-like statements but no laws and no procedures.

2. We are beginning to recognize that our older conceptions of method did not include the distinction between the interaction of the teacher with the content and his interaction with pupils. This distinction poses questions of major import. For example, how can rigorous handling of concepts, principles, procedures, and the like be maintained without turning off the pupils? Can it be done by improving the ingratiating techniques while more and more rigor is built into content strategies? Or can rigor and depth of insight resulting therefrom induce their own motivation?

3. Research has already uncovered certain variables--some related to content, e.g., clarity, and some related to the pupil, e.g., acceptance of pupil responses--which are related to achievement. These strengthen expectations that further progress is likely.

4. Teaching behavior has been analyzed into many kinds, most of which can be reduced to a few basic categories. This opens up the possibility for fruitful research in both the cognitive and affective aspects of teaching behavior.

Nevertheless, it must be noted that two basic aspects of teaching continue to be ignored by all of these studies: the diagnostic-remedial and the preventative dimensions of teaching behavior. No matter how well the various moves are performed and the strategies executed, no matter how skillful the teacher is in the techniques of ingratiation, some pupils will not learn. Except in reading and arithmetic, we have yet to study how teachers detect the learning difficulties of pupils and by what moves they attempt to cope with these problems. It seems reasonable to suppose that experienced teachers of science, social studies, literature, and so on, have learned or recognize some of the difficulties their pupils have in learning, and they have found a number of moves that help pupils overcome them. Information about what teachers already know and do would likely open up a number of fruitful lines of research.

Moreover, we know that some experienced teachers already know how to prevent certain difficulties in learning and in classroom conduct. But our knowledge of this aspect of teaching is meager. Some research in preventive pedagogy has been conducted, notably the studies of Kounin (1970) in classroom discipline and management. He has shown from analysis of classroom behavior that teachers who behave in certain ways, e.g., make smooth transitions from one activity to another, have fewer instances of disruptive behavior than teachers who do not do so. Work on preventative pedagogy is in its infancy, and further research could be fruitful.

References

Dewey, J. Democracy and education. New York: Macmillan, 1916.

Dewey, J. How we think. New York: Heath, 1910.

Flanders, N. A. Analyzing teaching behavior. Reading, Mass.: Addison-Wesley, 1970.

Henderson, K. A. Guide to clear thinking. Privately published, 1953.

Kilpatrick, W. H. Foundations of method. New York: Macmillan, 1925.

Kounin, J. A. Discipline and group management in classrooms. New York: Holt, Rinehart & Winston, 1970.

Morrison, H. C. Practice of teaching in the secondary school. Chicago: University of Chicago Press, 1926.

Morrison, H. C. Basic principles in education. Boston: Houghton Mifflin, 1934.

Ryans, D. G. Characteristics of teachers. Washington: American Council on Education, 1960.

Rosenshine, B. & Furst, N. Research on teacher performance criteria. In B.O. Smith (Ed.), Research in teacher education. Englewood Cliffs, N. J.: Prentice-Hall, 1971.

Skinner, B. F. The technology of teaching. New York: Appleton-Century-Crofts, 1968.

Smith, B. O., Meux, M. O., Coombs, J., Nuthall, G., & Precians, R. A study of the strategies of teaching. Urbana: Bureau of Educational Research, College of Education, University of Illinois, 1967.

Smith, B. O., & Meux, M. O. A study of the logic of teaching. Urbana: University of Illinois Press, 1970.

Thorndike, E. L. Darwin's contributions to psychology. University of California Chronicle, 1909, 12 (1), 65-80.

Thorndike, E. L. Educational psychology. New York: Teachers College, 1914.

Thorndike, E. L. Animal intelligence. New York: Hafner, 1965.

Toward the Development of Pedagogical Theory in Mathematics

Kenneth B. Henderson

Professor Emeritus

University of Illinois--Urbana-Champaign

This paper is based on two assumptions that I would like to make explicit at the outset. One is that it is possible to develop pedagogical theory in mathematics. I state this explicitly, for there are some individuals concerned with mathematics education who seem to doubt that this is possible. They appear to believe that teaching is a skill which, though complex, is learned by performing the act. The model is the apprenticeship one. The teacher-trainee serves as an apprentice to a master teacher who demonstrates effective teaching and serves as a model. When the teacher-trainee teaches, he is observed by the master teacher who offers suggestions for improvement. But, there is no structure to the suggestions, no theoretical basis, and only a meager rationale.

Only a few years ago at a conference on research in mathematics education Robert Davis (1967a) expressed doubts about the possibility of developing pedagogical theory in mathematics. To Davis, teaching is a skill or an art not amenable to description, analysis, and generalization.

It is well known that many members of departments of mathematics not only doubt that there is pedagogical theory in mathematics, but also doubt that it is possible to develop such theory. In contradiction to this theoretical position, I prefer to hypothesize that it is possible to develop this kind of theory.

A second assumption is that efforts to the end of developing such theory are worthwhile and deserve to be encouraged. To justify this assumption, I turn to the contributions of a theory--any theory--that disposes individuals interested in development of theory to devote their intellect, time, and efforts to this end.

Contributions of a Theory

The most obvious and dramatic contribution of a theory is that it enables those who understand it to adjust their behavior so as to obtain, in so far as possible, what they want. Sound pedagogical theory enables a teacher to predict what will occur under certain given conditions. The teacher then has some control over the situation; he does not have to operate only on trial and error. Theory is particularly useful when a teacher faces a problem he has not faced before, and his old behavior patterns cannot cope with it.

I am reminded of a story by William Everitt, former dean of the College of Engineering of the University of Illinois. Dean Everitt said that an engineer is an individual who can tell before a truck crosses a bridge whether the bridge will collapse. After the truck has crossed, any damn fool can tell. Because of his knowledge of theory, the engineer can predict with confidence. Then he can advise the truck driver whether to cross the bridge and the road commissioners what weight restrictions to set.

A second contribution of a theory is its explanatory power. A characteristic of man is his intellectual curiosity; once aware of certain states of affairs, he wants to know why they are as they are. Why can individuals who do not understand a particular algorithm, e.g., the long division algorithm, still employ it and get correct answers? (They comprehend it in the sense that they know what to do but do not understand why it "works.") An explanation of this phenomenon would consist of supplying facts and/or generalizations derived from psychological theory from which the phenomenon can be inferred; in other words, to subsume the phenomenon under one or more generalizations in psychological theory. Once this is done, we feel we understand; we know why individuals who do not understand a particular algorithm can still use it and get correct answers.

Now every theory, sound or unsound, enables explanation. What distinguishes sound from unsound explanation of phenomena is that the reasons proffered are known to be true, are not ad hoc, i.e., they explain phenomena other than the one being considered and are not tautologies. Sound theory promotes confidence in one's ability to cope with situations he experiences.

A third contribution of a theory is its use in generating researchable hypotheses. We know that some mathematics teachers who, on occasions teach by inductive discovery find that some students do not make the discovery at all and others verbalize one for which there are counterexamples. Why do such students not arrive at the correct conclusion which usually is a generalization? Perhaps because the teachers in choosing instances do not sample adequately from the domain over which the generalization to be discovered holds. Perhaps because they do not sequence the instances so as to facilitate the abstraction of the pattern. These are hypothetical explanations, i.e., they conflict with no principles in the theory, but we do not know whether they are true. We now hypothesize that if the teachers were taught to sample adequately and how to sequence instances to make the apprehension of a pattern--similarities amidst differences--

easy, and are given the rationale for doing these things, fewer students would experience the difficulties identified. This hypothesis is amenable to testing by experimentation. Whatever the outcome, we have a proposition to add to pedagogical theory. A good theory enhances its own advancement.

You will note that I have chosen to defend a theory on utilitarian grounds. I did this because in our society such an argument is persuasive. But I would not care to maintain that the utilitarian aspect of a theory is the only ground for accepting it. To some, a contribution of a theory is the organization it provides of the knowledge of the field over which the theory holds. People who value order, the portrayal of relationships, and deductive power as ends in themselves find satisfaction in the study of theory and in its development.

You may remember that I began with an identification of two assumptions on which this paper is based. I now state what I believe to be a fact. It is that there now exists pedagogical theory pertaining to mathematics. Let us turn to how this theory has been developed in the past and then to an analysis of the theory.

Approaches to the Development of Pedagogical Theory

One approach to the development of pedagogical theory has been for various teachers to reflect on their own experience and extract from it suggestions for how to do something. Thus, in mathematics education we see articles in professional journals and sections in textbooks on methods on how to teach that the product of two negative numbers is a positive number, how to show that the empty set is a subset of every set, how to teach students to prove theorems in geometry, how to teach the solution of equations, and how to teach students to solve worded problems, among other "how-to-do-its."

The suggestions distilled from experience can be at a higher order of generality, that is, they hold over a more extensive domain. For example, consider the following:

1. Make your assignments definite and clear.

2. Provide motivation.

3. Distribute your questions among the students in your class.

4. When teaching by guided discovery, make the students test their conjectures; don't provide them extraneous cues by voice, gestures, or facial expressions.

Usually, a rationale for the prescriptions can be given. This may be a deductive argument based on propositions believed to be true, or it may be by supplying evidence that following the prescriptions will attain some desirable end.

I do not contend that this approach has no merit. It is the case that mathematics teachers value how-to-do-it articles and speeches. These are practical and not couched in language replete with fuzzy pedagogical concepts. Yet I do not see this approach leading to substantial theory in mathematics education. Generalizing from personal experience leads to propositions that are idiosyncratic rather than nomothetic. The set of prescriptions emanating from this approach are unsystematic and lack structure. The prescriptions have restricted generality and hence restricted explanatory power. Moreover, they generally are not productive of researchable hypotheses which advance the theory. The approach tends not to be discriminating because it usually is not based on much analysis.

A second approach to the development of pedagogical theory is regarded as more scientific and valid. It presumes that there are implications for teaching from more basic theories, e.g., learning theory, social psychology, sociology, logic, or communication theory. For example, from learning theory, we presume that we should be able to deduce prescriptions directing how we should teach.

There are, or should be, misgivings associated with inferring pedagogical principles from more basic theories, e.g., learning theory. One misgiving stems from the evidence on which the principles are grounded. Typically, principles of learning are asserted whose variables seem to be universally quantified--all learners, all learnable objects, all situations. When one assesses the experimental evidence purporting to substantiate these propositions, he is beset with the feeling that such extrapolation is not warranted. Much of what we know about learning has come from the laboratory rather than from the classroom. Many of the propositions in learning theories are best supported by evidence on how nonhuman animals learn, e.g., rats, cats, dogs, and monkeys. These cannot be influenced by language as humans can, yet language is the chief resource of the mathematics teacher. When I read some of the unrestricted prescriptions about how to teach that are presumably based on learning theory, I think of the statement by the old cracker-barrel philosopher, Josh Billings, "It's better to know nothing than to know what ain't so."

The second misgiving rests on pragmatic grounds. It may well be that this approach enables us to explain some pedagogical phenomena, and this is better than nothing. But it is not distinguished by its ability to yield confirmed predictions, and this is what the practitioner expects of a theory. A teacher frustrated by the disparity between what educational psychology is alleged to be able to do and what, in fact, it can do, characterized an educational psychologist as a person who, given the facts in a pedagogical context, can predict what will happen, and then when his predictions do not eventuate can explain why they did not.

Even some educational psychologists have misgivings. Bugelski (1964) in his book, The Psychology of Learning and Educational Practice, says the following:

> The educational enterprise is a vast and complex one, involving as it does the training of people from kindergarten through graduate and professional schools in a wide variety of skills and knowledge. To apply the psychology of learning in its present state to such a tremendous field of activity should give even the boldest psychologist some pause. (p. 14)

Ernest Hilgard (cited in Bugelski, 1964) says "There are no laws of learning that can be taught with confidence" (p. 14). Kenneth Spence (cited in Bugelski, 1964), also a psychologist, echoes Bugelski's doubts.

> No definite answer, of course, can be given at the present time, for as yet none of them (i.e., learning theories) is sufficiently abstract or complete to account even for all the laboratory findings. (p. 15)

There are evidently grounds for Bugelski's assertion, "The teacher might well be wary of anyone who suggests some change on the basis of his knowledge of learning psychology" (p. 15).

An Alternative Approach

Instead of attempting to derive pedagogical theory from learning theory, one might study teaching as a phenomenon per se. The observations and records, e.g., audio- and videotapes, can be analyzed either in terms of the kind of students or the kind of subject matter taught or the interaction of these. Pedagogical models can be developed that describe how the teacher teaches. These then can be tested experimentally to see how well they explain and predict. Under this approach, pedagogy would emancipate itself from psychology. Incidentally, it was this very approach that enabled psychology to emancipate itself from philosophy.

Observations are always screened through a network of concepts and values. If mathematics teachers are observed--desirably in a classroom situation since most mathematics is taught in this manner--without certain concepts and values to guide the observations, the observer does not take certain things into account. Supplying fruitful concepts is a contribution of basic theories such as those mentioned above. The records of the observations can be analyzed in terms of kinds of students or kinds of teachable objects, e.g., concepts, facts, principles, skills, and values, or the interaction of these. Drawing on whatever basic theory seems relevant and fruitful, pedagogical models can be developed that describe how the mathematics teachers teach the various teachable objects. These, then, can be tested experimentally to see how well they explain why teachers teach as they do and predict how well students learn.

As some of you may know, this is the approach I and some of my graduate students have used. I have taken the position that the set of teachable objects, e.g., knowledge and beliefs, is not identical to the set of learnable objects. Nor is either a subset of the other. There are some items of knowledge that can be taught to but not learned by a particular group of students. And there are some objects that are learned but not taught, e.g., mathematical intuition and values that we characterize as being caught rather than taught. Basic to this point of view is the belief that it is not fruitful for either clear thinking or research to define teaching in terms of learning.

Which basic theories should be used to analyze the observations? The answer to this question depends partly on what the theoretician thinks is important and fruitful, and what he is willing to sacrifice. As I continue the analysis, you will be able to tell which theories I consider fruitful and which appear to get short-changed.

I doubt that it is profitable in the existing state of our theory to consider teaching in the abstract. The tendency to do so results in unrestricted generalizations or prescriptions. Practioners become aware of so many counterexamples they lose confidence in the theory. I believe that at present we should take as our object of a study a teacher teaching some kind of teachable object to one or more persons. For example, the mathematics teacher teaching various kinds of knowledge and beliefs about mathematics to his students. To me, this makes theory of knowledge relevant. There are different kinds of knowledge and different kinds of beliefs. There is knowledge that such is the case, and there is knowledge of how to do something. Although these are related, they are distinct. Neither is a sufficient condition for the other. Of knowledge of what is the case, some of it is empirical and some is analytic. In elementary school and junior high school, mathematics is taught as though it is empirical. As the student matures mathematically, hopefully it is taught as it really is, i.e., analytic knowledge.

We have different rules for assigning the truth-value "true" to these two different kinds of knowledge. Within each of these kinds of knowledge, there are concepts, singular statements, generalizations, and prescriptions. I would argue that the logic of each of these subsets is distinctive. And the distinctive logic has implications for teaching provided the teacher seeks to show the students how we know or accept the various items of knowledge.

Knowledge of how to do something can be classified in terms of whether it is based on empirical generalizations or an analytic generalizations. How to use a micrometer is based on empirical knowledge of the relation between one turn of the thimble and the resulting change in the graduations exposed on the barrel, what happens if the spindle is not tightened enough or tightened too much, and of the function of the ratchet. It is also based on knowledge of which objects are amenable to being measured by the micrometer and which are not. Knowing how to factor a polynomial over some set, e.g., the integers, is based on analytic knowledge, viz., the distributive principle of multiplication over addition.

There are different kinds of beliefs, e.g., about what mathematical knowledge and proficiency are necessary for certain jobs, about what ought to be done in solving a problem, about what is important in mathematics, and about the nature of mathematics. Like items of mathematical knowledge, beliefs about mathematics are teachable objects.

One can hypothesize that the kind of teachable object should make a difference in teaching just as the kind of behavior of the student should. For example, definitions have an arbitrariness associated with them that is not associated with generalizations. Saying that a definition is true or saying that it is false is not the same as making analogous statements about a factual statement. Surely a mathematics teacher should draw a distinction between a factual statement and a value judgment and hence teach them differently.

Since much of the teaching of mathematics is via language, semantics seems relevant and productive of insights. Logic seems relevant because usually the mathematics teacher appeals to the rational aspect of the student's personality. Logical connectors like because, therefore, if-then, and, or, and their cognates abound in classroom dialogues.

Moves and Strategies

In our analyses of classroom dialogues, we find teachers and students using language to define, describe, compare, contrast, instantiate, characterize, identify, assert, generalize, imply, infer, justify, direct, exhort, classify, exemplify, and others. (Someone who is familiar with John Austin's semantics will recognize these as illocutionary acts which also have perlocutionary potential.) We have made use of the concepts of a move and a strategy. A move is a bit of discourse in which language is used in a certain way such as those listed above. A strategy is a sequence of moves.

Our approach has been to audiotape classroom teaching and then analyze the transcriptions utilizing the theories mentioned above to identify the teaching of various objects, e.g., concepts, principles, and skills. Then each of these kinds of teaching is further analyzed in an attempt to identify moves (bits of monologue or dialogue) and strategies (sequences of the moves) which are evident. This part of the research is naturalistic, conceptual, and descriptive. But it is basic to subsequent experimentation leading to principles of teaching.

I shall not say much at this time about the moves and strategies in teaching a concept. I have written about this in a chapter in the Thirty-third Yearbook of the National Council of Teachers of Mathematics (Henderson, 1970), and, more recently, in a textbook on methods of teaching secondary school mathematics (Cooney, Davis, & Henderson, 1975). In both of these I offered a classification of moves, a simpler classification in the methods book than in the chapter in the Yearbook. In both I proposed a classification of concepts since it appears that the kind of concept taught, to some extent, determines the moves that can be used. For example, there are some concepts for which some moves are logically impossible. It is logically impossible to teach a nondenotative concept--by nondenotative concept I mean one like even number greater than two or the greatest upper bound of the integers--by giving examples. It is logically impossible to give a definition of any undefined term in mathematics. And, one cannot give a counterexample unless a false generalization has been asserted.

In the methods textbook there are chapters on the teaching of principles in which a classification of principles is offered and moves and strategies identified. In passing, I might mention that by using the concept of move and strategy it is possible to sharpen the distinction between expository teaching and guided discovery, and between deductive guided discovery and inductive guided discovery.

Once moves and strategies have been identified, it now becomes possible to ascertain under what conditions various strategies are correlated with learning on the part of students. One can hypothesize that certain strategies will be effective for slow learners. If one wants to conceive of levels of learning like those in the Taxonomy of Educational Objectives by Bloom (1956), he can hypothesize that certain strategies will correlate with certain levels of learning. I would venture a guess that for certain kinds of concepts, e.g., precise vs. vague, complex vs. simple, abstract vs. concrete, disjunctive vs. conjunctive, certain strategies will be more effective than others. An analogous conjecture can be made about kinds of principles, e.g., complex vs. simple, prescriptions vs. generalizations. One can also conjecture that among all the strategies some will prove more efficient than others, that is, some will produce the same level of comprehension in fewer moves. For example, I doubt that inductive guided discovery is an efficient strategy for teaching a complex principle, e.g., a generalization conditional in a form in which several conditions are conjoined in the hypothesis. Hence, the approach of observing teachers teaching certain teachable objects to certain groups of students, analyzing the teaching, and developing models is productive of researchable hypotheses that are capable of empirical verification. Dossey (1972), Malo (1974), Rector (1966), Retzer (1967), Rollins (1966), and others have tested some of these.

I would feel more comfortable defending the approach to developing pedagogical theory in mathematics that I have just been talking about than the particular models that some of us have determined by using this approach. Any time one uses analysis, of necessity, he focuses on certain factors and chooses to neglect others. It may well be that other theoreticians who would make use of the approach I have suggested would come up with models that are more fruitful than those we have determined. Be that as it may, it is apparent that the specific pedagogical models I have described have been developed independently of learning. To anyone who understands theory construction, it is obvious that there cannot be an analytic connection, viz., by definition, between teaching and learning. It must be a contingent connection, e.g., if a teacher makes such and such moves, there is or is not, as a matter of fact, a correlation with learning on the part of the students he is teaching. The models we have developed enable such a contingent connection.

Possible Additional Research

So far, our research has concentrated on ventures in which just one concept, principle, or skill has been taught. Yet we know that in actual classroom situations usually a sequence of items of knowledge or belief is taught. The teaching of a concept may be embedded in the teaching of a principle; or a principle may be taught once a concept whose grasp is necessary for comprehension of the principle has been taught; or the teaching of a principle may be embedded in the teaching of a skill.

Pavelka (1974), using the approach I have described, identified what she denoted as modes of teaching two or more concepts. One was the consecutive mode. Suppose two concepts A and B are taught in a temporal sequence. In the consecutive mode, the last move in teaching A precedes the first move in teaching B. We recognize this mode as prevalent in textbooks where there is no feedback from the learner to the teacher and no diagnosis by the teacher is made. Pavelka found that this was the mode most frequently used by mathematics teachers.

A second mode is embedive. In this mode the first move in teaching concept B follows the first move in teaching concept A and the last move in teaching B precedes the last move in teaching A. In other words, all the moves in teaching B are between the first and last moves of teaching A.

A third mode is overlapping. In this mode the first move in teaching A precedes the first move in teaching B and the last move in teaching A precedes the last move in teaching B. For example, the teaching of a concept of a mixed number may overlap the teaching of a concept of an improper fraction, and both of these may overlap the teaching of a concept of a proper fraction.

What modes would be found when more than two concepts are taught in one class period? Is there a high correlation between the diagnostic ability of a mathematics teacher and the use of the embedive and overlapping modes? Do the same modes appear when a sequence of principles is taught? What modes would be found in a sequence involving both concepts and principles? All these questions seem interesting to me.

So far we have not studied the teaching of values in mathematics. I postponed this because it seemed easier to study the teaching of concepts, principles, and skills. Could the concepts of moves and strategies be used to analyze the teaching of values? If so, what moves would be found? I can speculate, but I do not know.

Finally, I have been interested in how students learn concepts without being taught them. There are many concepts in this category. Suppose students have been taught a concept of equilateral figures and a concept of a triangle. Thereafter, the teacher would probably use the term equilateral triangle without deliberately teaching the concept. How do students acquire this latter concept? I can advance an explanation based on logic. But I do not know whether this is the way the student's mind operates. Or suppose students have been taught a concept of an equilateral triangle. The teacher might subsequently use the term equilateral quadrilateral without taking time to teach this concept. If the students acquire the latter concept simply by hearing the term used, what is their reasoning process? Again, I can offer a hypothetical explanation, but I do not know whether it is true of any student or of many students. Perhaps if we were able to answer questions like these, we would be better able to tell under what conditions a mathematics teacher can safely use a term without taking time to teach the concept it designates.

Elements in a Pedagogical Theory

Regardless of what approach is used to develop pedagogical theory, I would expect to find certain elements in the theory. Earlier, I suggested that a theory is a network of concepts and statements which are either knowledge or beliefs. In an empirical theory, I would expect some of the concepts to be of sensed objects. Thus, in pedagogical theory we have concepts of students, teachers, classroom dialogue, overhead projectors, and colored chalk, among many others because we sense each of these entities.

Other concepts are inferred entities--theoretical constructs. We do not sense these. We invent rather than sense such concepts as motive, aptitude, interest, proficiency, appreciation, understanding, concept, principle, skill, attitude, and many others because they aid us in explanation and in identifying researchable hypotheses. We explicate such concepts, if we ever do explicate them, in terms of sensed concepts and facts. I cite the current interest in behavioral objectives as an attempt to explicate certain concepts which are inferred entities (comprehend, understand, appreciate) in terms of observable entities (state, write, draw, list, choose).

The statements in a theory, unlike the concepts, have the characteristic of being known to be or believed to be true. Since pedagogical theory is an empirical theory, I would expect most of the statements to be empirical or factual. Yet I would expect to find some analytic statements. Here are some statements that seem to me to be analytic:

1. Given any test of computation in arithmetic, fifty percent of the students who take the test will not score higher than the median.

2. If a teacher has induced a desire to learn in a student, he has motivated the student.

3. If an objective is unachievable in the time available for schoolwork, the teacher will not be successful in helping students attain it.

4. Experience shows that students will either enjoy mathematics or not enjoy it. (Experience does not show this; logic does.)

5. If what a superintendent says is not credible, it will not be believed.

6. Relative to other people, a person's adjustment mechanisms will be toward other people, away from them, or against them.

In quoting these statements, which I found in books pertaining to teaching, I am not ridiculing the writers. Analytic statements are unavoidable in a theory. They are statements in the object language which are immediate implications of definitions or are instances of tautological formulas in logic.

What is necessary for both a theoretician and a practitioner is to be able to tell when a sentence is used to make an empirical (factual) statement and when it is used to make an analytic statement. Let us take some examples.

"Slow learners cannot handle abstractions as well as average learners." Is this statement used to make a factual or an analytic statement? One cannot tell without obtaining answers to one or more questions. How are slow learners identified? If they are identified by tests whose items test the ability to handle abstractions, the statement is analytic. On the other hand, if they are identified by behavior other than their ability to handle abstractions, the statement is factual.

"Good teachers motivate their students." Is it conceivable, i.e., possible, that a teacher could not motivate his students and still be regarded as a good teacher? If the answer is "yes," the statement is empirical. Whether such teachers exist as a matter of fact is irrelevant; all that matters is that it is logically possible for such teachers to exist. If the answer is "no," the statement is analytic. Motivating students is a logically necessary condition for being a good teacher.

"Students never fail; only teachers fail." Even though this is declarative in form, I doubt that it is used to make a statement. It is a slogan. Nevertheless, there are individuals who think that it makes a statement. Such individuals need to be asked to describe an experiment which would test this statement. If they cannot do this (and I think it is impossible to do so), they are using the statement as analytic. This they have every right to do. It is akin to "If the student hasn't learned, the teacher hasn't taught." which is one of the basic tenets in their theory of teaching. Authors who make such an assumption should not talk or write as though it were a matter of fact. (As an aside, perhaps one of the desirable outcomes of the thrust of accountability will be for teachers to divest themselves of these unfortunate slogans.)

Some of the confusion in educational theory stems from the unawareness of the use of sentences which are used to make factual statements and which are used to explicate the semantics of the theory. When parents, teachers, or theoreticians disagree about a statement, the first thing to do is to ascertain how the statement is being used. Once this is settled, the kind of evidence that is relevant is determined. The search is for facts if the statement is used to make a factual assertion; it is for definitions or logical propositions if the statement is used to make an analytic assertion.

I now turn to another element that is present in a theory like pedagogy which is regarded as primarily practical. This element is value judgments. In well established and objective theories, the value judgments are implicit. Thus, nowhere in mathematical theory is the explicit judgment that abstractness is desirable. This goes without saying. Yet in pedagogical theory we find abundant use of rating terms, e.g., good, important, significant, desirable, worthwhile, and their antonyms, to express value judgments. Their prevalence may be explained by the lack of agreement on basic values in education. With lack of agreement, values need to be made explicit, for they, in part, determine choices of action.

I presume that I need not argue that value judgments have a logic that is distinct from that of analytic and factual statements. Only the dogmatic or omnicient individual would attempt to extend Tarski's paradigm

'Snow is white' is true if and only if snow is white

to

'Mathematics is good' is true if and only if mathematics is good.

The final element in a theory I shall discuss appears only in a theory which is regarded as primarily practicable. This element is hortatory statements, i.e., exhortations to do so and so or prescriptions concerning how to do something. Examples are:

1. A teacher should meet the needs of his students.

2. Subject matter that is too difficult for the students in a class should not be chosen.

3. Give the students experiences in which they have to apply what they haved learned.

4. Make students check their answers to problems.

These exhortations are inferences from factual statements and values. They are used to advise, urge, or direct behavior. Hence their logic is different from that of assertions. Rather than being judged on whether they are correct descriptions of the facts--Tarski's paradigm--they are judged either in terms of whether they can be defended by facts and values, or by whether following them attains some end that presumably is accepted by both the person who utters the hortatory statement and the person who accepts it. We might regard the generalizations in a theory, both analytic and empirical, and the hortatory statements as the principles in the theory. It is these that give the theory power.

Why are there more hortatory statements in pedagogical theory than in physics? It is not because pedagogy is practical and physics is theoretical or that pedagogy is primarily concerned with human behavior and physics is not. I suggest it is because the generalizations in pedagogical theory are less well established and the values less well accepted. If I wish to cool my cup of coffee, there are clear implications from established generalizations in physics which tell me what to do. I don't need prescriptions. But if I want to teach mathematics to certain groups of students, e.g., those who live in the inner city, there are no generalizations whose variables are quantified over well-defined domains from which I can infer reliable prescriptions. I need to be offered advice in hortatory language. As pedagogical theory attains generalizations for which there are few, if any, counterexamples, the need for hortatory language will disappear. Such generalizations will be products of theory construction and experimental research. The stimulation of both of these is, I judge, the main thrust of this conference.

Summary

In case I lost you somewhere in this lengthy analysis, let me summarize. I contended that it is both possible and desirable to develop pedagogical theory in mathematics. It can be developed by many teachers on their own teaching and collating the judgments, by deduction from more basic theories, or by systematic observation and analysis of what goes on when teachers teach mathematics to groups of students. I advocated the latter and pointed out theories that seem relevant to the analysis. I deliberately ignored learning theory, for this discussion is on teaching and I contend that teaching needs to be identified (defined) independently from learning if we are to avoid an analytic connection between the two. Analysis of the teaching of sets of pervasive objects like concepts, principles, skills, and values will provide models which can be tested to ascertain the conditions under which they are correlated with learning on the part of students. Such an approach will produce a dependable theory. This theory will be composed of concepts, sensed or inferred; principles, empirical or analytic; and value judgments. Those principles which are hortatory statements facilitate choices for those who do not have much command of the other elements in the theory.

References

Bloom, B. S. (Ed.). Taxonomy of educational objectives. New York: David McKay, 1956.

Bugelski, B. R. The psychology of learning applied to teaching. Indianapolis: Bobbs-Merrill, 1964.

Cooney, T. J., Davis, E. J., & Henderson, K. B. Dynamics of teaching secondary school mathematics. Boston: Houghton Mifflin, 1975.

Davis, R. B. Mathematics teaching--with special reference to epistemological problems. Paper presented at the National Conference on Needed Research in Mathematics Education, Athens, Georgia, September 1967. (a) (A summary of this paper can be found in Davis (1967b).)

Davis, R. B. The range of rhetoric, scale, and other variables. Journal of Research and Development, 1967, 1 (1), 51-74. (b)

Dossey, J. A. The relative effectiveness of four strategies for teaching disjunctive concepts in mathematics (Doctoral dissertation, University of Illinois, 1971). Dissertation Abstracts International, 1972, 32A, 5467-5468. (University Microfilms No. 72-12,142)

Henderson, K. B. Concepts. In M. F. Rosskopf (Ed.), The teaching of secondary school mathematics. Thirty-third Yearbook of the National Council of Teachers of Mathematics. Washington, D.C.: The Council, 1970.

Malo, G. E. Differential treatments in learning disjunctive concepts in mathematics (Doctoral dissertation, University of Illinois, 1974). Dissertation Abstracts International, 1975, 35A, 4051. (University Microfilms No. 75-363)

Pavelka, E. B. Moves, strategies, and models of a concept venture in secondary school mathematics (Doctoral dissertation, The University of Illinois, 1974). Dissertation Abstracts International, 1975, 35A, 7183. (University Microfilms No. 75-11,667)

Rector, R. E. The relative effectiveness of four strategies for teaching mathematical concepts (Doctoral dissertation, University of Illinois, 1968). Dissertation Abstracts, 1968, 29A, 520. (University Microfilms No. 68-12,184)

Retzer, K. A. The effect of teaching certain concepts of logic on verbalization of discovered mathematical generalizations (Doctoral dissertation, University of Illinois, 1967). Dissertation Abstracts, 1967, 28A, 1351-1352. (University Microfilms No. 67-11,909)

Rollins, J. H. A comparison of three stratagems for teaching mathematical concepts and generalizations by guided discovery (Doctoral dissertation, University of Illinois, 1966). Dissertation Abstracts, 1966, 27A, 711-712. (University Microfilms No. 66-7802)

The Role of Relative Efficacy Studies in the Development of Mathematical Concept Teaching Strategies: Some Findings and Some Directions

John A. Dossey

Illinois State University

The Search for a Model of Concept Teaching

The Need for a Model

For several years the subject matter of methods courses for prospective teachers of mathematics has been built around a core of efficacious prescriptions drawn from teachers' past experiences, educational psychology, theories of learning, and mathematics. While each of these areas has something to offer the teacher, they do not provide the mathematics instructor with a firm foundation for educational decision making concerning the teaching process.

These sources do not provide a rational framework for a careful analysis of the teaching act. The field of prior experiences is fixed in time, and the conditions which led to a particular happening may never eventuate again. The findings of educational psychology and the ramifications of theories of learning do not offer the insight which was once expected of them; rather they focus on the learning aspect of the classroom situation and not on the teaching act. Hence, they are descriptive and not predictive. They do not provide a mirrored surface from which we can deduce theories of teaching. Other shortcomings of the findings in these areas and their values to the development of a theory of teaching are discussed in a paper by Henderson (1972). The field of mathematics does lend some guidance in that its structure provides some information concerning the ordering of concepts and generalizations. However, it does not aid in the selection of teaching strategies or materials for classroom instruction.

These shortcomings have prompted several educators to call for the development of a "theory of instruction" or a "theory of teaching." Gage (1963) feels that such a theory should

> attempt to explain how it is that the behavior of one person, a teacher, can influence the behavior or learning of another person, a student. This kind of theory would attempt to explain instances of teaching, i.e., of interpersonal influence resulting in learning. (p. 134)

Bruner (1966) builds on Gage's call for a "theory of instruction" by stating four conditions such a theory should fulfill:

> First, a theory of instruction should specify the ways which most effectively implant in the individual a predisposition toward learning--learning in general or a particular type of learning. . .
>
> Second, a theory of instruction must specify the ways in which a body of knowledge should be structured so that it can be most readily grasped by the learner. . .
>
> Third, a theory of instruction should specify the most effective sequences in which to present the materials to be learned. . .
>
> Finally, a theory of instruction should specify the nature and pacing of rewards and punishments in the process of learning and teaching. (pp. 40-41)

While no such universal theory exists, several different models have been proposed to serve as a framework for the analysis and study of teaching. Nuthall and Snook (1973) have provided the educational community with an overview of several of the more popular models for teaching. Their analyses consist of a study of the origins of the models, research stimulated by the models, and issues generated by playing the models off against each other.

These models have been responsible for the generation of a multitude of studies which have tested hypotheses concerned with attempts to better understand the act of teaching and to develop better methods of instruction. This paper will consider the development of the rational model of teaching as it interacts with concept teaching in mathematics. The origins of the Smith-Henderson model for concept teaching will be analyzed, the research generated by it reviewed, and a model for further studies will be proposed.

Smith's Model

The rational model for the study of teaching has its origin in the work of B. O. Smith and his associates at the University of Illinois. Their work on the development of a model for teaching began in the late 1950's. Smith (1956) argued that teaching is primarily a verbal activity.

With this in mind, the first attempts in the building of the model for teaching were audio-recordings of the interactions which took place in classrooms when "teaching" supposedly was taking place. The transcripts of the resulting dialogues were then subjected to a rigorous logical analysis to determine the manners in which teachers use language in the classroom.

This analysis consisted of factoring the total classroom discourse into units known as episodes, which consisted of a series of completed verbal exchanges between two or more speakers, and monologues, which consisted of the words of a single speaker. The examination of the various uses of language revealed that teachers use language to define, designate, classify, explain, compare-contrast, evaluate, and offer opinions (Smith, Meux, Coombs, Eierdam, & Szoke, 1970). This study was quickly followed by a second study. Here, the emphasis was shifted from the factoring of the total discourse into recognizable units to a study of the sequencing of these units into strategies for dealing with different types of subject matter material, such as concepts and principles (Smith, Meux, Coombs, Nuthall, & Precians, 1967).

Smith characterizes the action in the classroom as a game played between a teacher and a class. This situation involves both mutual and conflicting goals. The teacher attempts to attain his goals through his action in the classroom, whether it is with student cooperation or student resistance. It is here that the game theory concept of a strategy enters. Smith et al. (1967) stated that:

> Pedagogically, strategy refers to a set of verbal actions that serve to attain certain results and guard against others. From a general standpoint, strategies may serve to induce students to engage in a verbal exchange, to insure that certain points in the discourse will be made clear, and to reduce the number of irrelevant or wrong responses as the students participate in discussion and so on. Of course, strategies also enhance the possibility that the cognitive import of the venture will be attained; that is to say, the objectives such as explications of concepts, elaborations of causal conditions, and the presentation of information will be successfully carried out. (pp. 49-50)

This formulation of a pedagogical strategy then served to provide a framework for a logical development of a model of classroom teaching in terms of teachers' verbal moves. The total verbal discourse of the classroom was divided into units aligned with the teacher's objectives. The venture unit contains all of the material related to a single overarching objective. Each venture identified in the discourse was then split into smaller divisions called moves. The moves are the various forms in which a teacher logically structures the subject matter material as he attempts to move toward his objective. It is the sequencing, or patterning, of these moves that gives rise to the concept of a strategy which Smith refers to in the passage quoted above.

One of the classes of ventures identified in this study of classroom teaching was that of a conceptual venture. The objective of ventures of this type was the development of a student's ability to tie a set of meanings and a particular term together in such a way that the student has command over usage of the term. This model of the relation of conceptualization has been expanded upon by Henderson (1963).

The analysis of the audiotapes of classroom verbal interactions showed that when teachers are dealing with the teaching of concepts, they are not concerned with teaching how to categorize or discover the defining attributes of a particular concept, as they are often called on to do in laboratory studies of concept attainment. Rather, Smith et al. (1967) found that the teachers were

> concerned with the explication of the kind of information which results in students being able to describe the concept, identify differences between the concepts and some other concept, and to understand the concept in learning about more advanced subject matter. (p. 60)

These findings were also noted by Carroll (1964).

When the taped classroom interactions were viewed from this point, Smith and his coworkers were able to detect four main categories of moves used by teachers in dealing with concepts.

1. Descriptive moves. In these segments of verbal discourse, the teacher gave characteristics of the concept of concern, analyzed the concept, or classified the concept.

2. Comparative moves. These moves consisted mainly of attempts to give analogies between concepts or to differentiate between concepts.

3. Instantiation moves. The moves in this set were concerned with providing examples, nonexamples, or the production of justification of either examples or nonexamples of the concept in question.

4. Usage moves. The only move identified in this class was the move called the meta-distinction move. The meta-distinction move is one in which the speaker discusses the term which names the concept of interest.

The four categories and the moves contained in each resulted in the following taxonomy of concept teaching moves:

I. Descriptive Moves

  A. Characteristic move
  B. Sufficient condition move
  C. Classification move
  D. Classificatory description move
  E. Relations among characteristics move
  F. Analysis move

II. Comparative Moves

  A. Analogy move
  B. Differentiation move
  C. Instance comparison move

III. Instantial Moves

  A. Positive instance move
  B. Instance enumeration move
  C. Negative instance move
  D. Instance production move
  E. Instance substantiation move

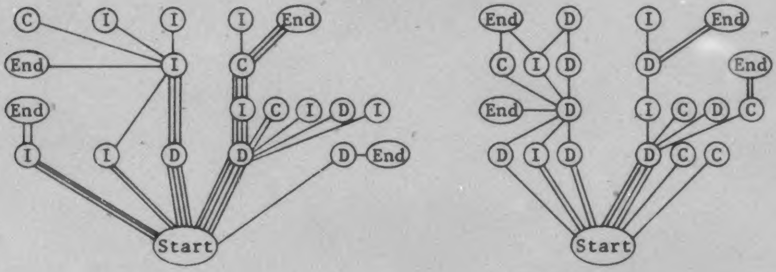
IV. Usage Moves

  A. Meta-distinction move

After identifying the classes of moves and the types of moves contained within them, Smith and his coworkers described the strategies they found in their analyses of audio interactions concerned with concept teaching. Seven patterns of concept teaching emerged, but only four were employed in over ten percent of the concept ventures studied (Smith et al., 1967).

The most prevalent pattern was the one involving only descriptive moves. This strategy was used in thirty percent of the ventures cited. The second most common strategy was that of a group of descriptive moves followed by a group of several instantial moves. This sequence occurred in twenty-seven percent of the ventures. The third type of sequencing pattern noted was that of a group of descriptive moves followed by a set of comparative moves. This strategy was used about sixteen percent of the time in concept venture situations. The fourth major strategy was found to be a combination of descriptive, comparative, and instantial moves. It made up about twelve percent of the concept ventures identified.

This same study (Smith, et al., 1967) also showed that the combination of moves used in teaching a concept, as well as their length and order, varied from one subject matter area to another. Some examples of strategies noted are shown in Figure 1. The numerals in the circles represent the type of move used with respect to Smith's taxonomy. The "D," "C," and "I," in the circles indicate whether the move was a descriptive, comparative, or instantial move. The multiple lines indicate the number of times the pattern of moves was observed. This analysis indicated that the logical nature of the concepts and their subject matter affiliations might determine both the nature of the strategy and the types of moves contained within it.



Chemistry Teacher

Biology Teacher

Figure 1. Some examples of teacher strategies for concept teaching (Smith et al., 1967).

Henderson's Model

A Study of the Strategies of Teaching (Smith et al., 1967), combined with Smith's earlier work, led Henderson to propose a set of moves for the teaching of concepts in mathematics (Henderson, 1967). This set of moves was designed with respect to a functional taxonomy of concepts which Henderson (1970) proposed to help in the selection of concept teaching moves and strategies for teaching concepts. This taxonomy will be discussed later in this paper.

Henderson noted that a term can be used in three different ways. First, a term can be used to talk about the characteristics of the elements in a term's referent set. This is said to be a use of the term's connotation. Second, a term is used in labeling objects or ideas as being members or nonmembers of the term's referent set. Here the teacher uses the term's denotation. In both of these instances, the teacher is talking about the concept in the object language. A third use of the term occurs in the metalanguage. Here the speaker uses the term in talking about the term itself. This type of usage of the term is said to use its implication. It was from these three uses of a term which names a concept that Henderson's taxonomy of concept moves was developed.

The following analysis gives a full explication of Henderson's taxonomy of concept teaching moves, as well as their relationship to the use of the concept naming term. The first moves in the taxonomy are the moves in the object language. These moves are divided into the set of characteristic moves, which deal with the connotation of the concept naming term, and the set of exemplification moves, which deal with the denotation of the concept naming term. The second set of concept teaching moves are the moves in the metalanguage. These moves, called definitional moves, deal with the implication of the concept naming term.

Henderson's set of moves used in the teaching of mathematical concepts is briefly stated in the following outline: (See Henderson, 1967 for a more detailed description).

I. Moves in the object language.

   A. Based on characterization. Characterization moves are those in which a person talks about the characteristics or properties of the objects in the referent set.

      1. Single characteristic
      2. Sufficient condition
      3. Necessary condition
      4. Classification
      5. Identification
      6. Analysis
      7. Analogy
      8. Differentiation
      9. Comparison and/or contrast of members of the referent set

   B. Based on exemplification. Exemplification moves are those in which a person names members and nonmembers of the referent set. Or one person designates an object and asks another to determine whether or not the object is a member of the referent set.

      1. Example
      2. Nonexample
      3. Counterexample
      4. Specification
      5. Exemplification accompanied by justification
      6. Nonexemplification accompanied by justification

II. Moves in the metalanguage

A. Stipulated definition. A meaning is ascribed to a term which is to designate a concept.

B. Reported definition. The conventional meaning of a term designating a concept is reported.

With Henderson's analysis of the various types of moves used in teaching a concept, one has a model for considering the relative efficacy of various strategies used by teachers and teaching materials in structuring conceptual ventures in mathematics. Expanded information on several of the above moves is available in a slightly modified version of Henderson's taxonomy in Dynamics of Teaching Secondary School Mathematics (Cooney, Davis, & Henderson, 1975).

The formulation of a classroom concept teaching strategy must take into account the logical nature of the concept, the nature and amount of information carried in various exemplars of the concept, and the theoretical advantages one type of move might possess over another type of move. The starting point for such studies might be in the translation of Smith's strategies into the nomenclature of Henderson's taxonomy.

Because Henderson's taxonomy has only two classes of moves in the object language, shifting from Smith's model to Henderson's model for teaching concepts condenses the strategies mentioned earlier into two general types. The strategy types noted earlier are now characterized as consisting entirely of characterization moves or as being a mixture of characterization and exemplification moves. If such sequences are given an alphabetic designation by letting a "C" denote a string of one or more characterization moves and an "E" denote a string of one or more exemplification moves, the strategies might be called C strategies, or CE, EC, CEC, or ECE strategies.

An analysis of the common concept strategies used by teachers in A Study of the Strategies of Teaching (Smith, et al., 1967) in terms of Henderson's "C" and "E" moves gives sequences as shown in Figure 2. The C' and E's in the circles indicate whether the moves are from the characterization set or the exemplification set. The set of letters at the end of the branches of the tree network indicate the type of strategy represented by the path through the network terminating at that mode.

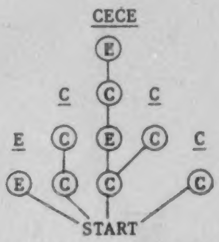


Figure 2. Examples of common strategies seen over many teachers.

Early Research Studies Concerning Concept Teaching Strategies

Ginther's Study

The next study which attempted to unravel the question of what strategies are used in teaching mathematical concepts was conducted by Ginther (1964). (Ginther's study is reported in Ginther and Henderson (1966).) Ginther made a survey of twenty-three algebra and geometry texts to identify the instructional strategies employed in the handling of definitions. Ginther classified the definition teaching moves on the basis of whether the moves designated the objects denoted by the term being defined, gave a set of necessary and sufficient conditions for an object to be labeled by the term being defined, or by giving a term claimed to mean the same thing as the term being defined. These three classes of definitions were called denotative, connotative, and synonymical definitions respectively.

The results showed that in both algebra and geometry connotative definitions were used most, and synonymical definitions were used least. An analysis of the percentages of each type used in each subject showed a difference between the types of definitions, as shown in Table 1 below (Ginther & Henderson, 1966).

Table 1

Percentage of Each Definition Type Appearing

In Algebra and Geometry Textbooks Surveyed

Definition Type	% of Algebra Moves	% of Geometry Moves
Denotative	29	10
Connative	66	89
Synonymical	5	2

An analysis of these percentages indicates that the authors of the texts surveyed may think that characteristics, or properties, of the objects in the concept's referent set are the most important factor in the attainment of a concept from written materials. This conviction seems to be stronger in geometry than in algebra, as the percentage of connotative definitions in geometry was 89, while the percentage of connotative definitions in algebra was 66. The denotative definition moves ranked second in order of usage in both the algebra and geometry materials. They accounted for 29% of the definitional moves in algebra and 10% of those in geometry. This finding might be interpreted as meaning that authors of textual materials in mathematics feel that the use of examples is more important in the teaching of algebra than it is in the teaching of geometry. Another interpretation for the drop in denotative moves as one changes from algebra to geometry could be the amount of page space required for denotative moves in a geometry text. The use of synonymical definitions was found to be the least used strategy in both subject matter areas. This finding agreed with Smith's et al. (1967) findings concerning the meta-distinction move.

When the total strategies for presenting definitions were examined, Ginther found four basic patterns. These patterns can be considered concept teaching patterns, for the definition teaching strategies were designed to place a specific meaning with a particular term. The basic patterns denoted by Ginther were explanation-definition, definition-explanation, explanation-definition-explanation, and definition.

Ginther's work was important in that it established the existence of four different instructional strategies which appeared in contemporary mathematics texts. In addition, it noted the differences in the use of the three classes of definitional moves in algebra and geometry texts.

Rector's Study

The first study to empirically test the relative efficacy of different concept teaching strategies for mathematical concepts was done by Rector (1968) (Rector's study is reported in Rector and Henderson (1970).) Rector arrived at four strategies to examine by making a liberal translation of Ginther's definitional strategies into Henderson's concept move language. This was done by substituting characterization moves for Ginther's definitional moves and exemplification moves for Ginther's explanation moves. This resulted in the following four strategies:

1. a set of exemplification moves followed by a set of characterization moves,

2. a set of characterization moves followed by a set of exemplification moves,

3. a set of exemplification moves followed by a set of characterization moves followed by still another set of exemplification moves, and

4. a set of characterization moves.

Using the alphabetic mode of representing strategies, these four strategies would appear as EC, CE, ECE, and C strategies respectively.

Rector compared the relative efficacy of these four strategies for their ability to influence student acquisition of eleven concepts from elementary probability theory. Rector developed programmed instructional materials in which the strategies, each consisting of five moves per concept, were structured as follows:

1. Characterization Strategy (C). This instructional strategy consists of five characterization moves.

2. Characterization-Exemplification Strategy (CE). This instructional strategy consists of a characterization move followed by four exemplification moves.

3. Exemplification-Characterization Strategy (EC). This instructional strategy consists of four exemplification moves followed by a characterization move.

4. Exemplification-Characterization-Exemplification Strategy (ECE). This instructional strategy consists of two exemplification moves followed by a characterization move followed by yet another set of two exemplification moves.

To test the relative effectiveness of the four strategies, Rector established a taxonomy of cognitive behaviors patterned after that developed by Bloom, Englehart, Furst, Hill, and Krathwohl (1956). Rector collapsed Bloom's taxonomy into the following three levels: (a) level one, knowledge and comprehension; (b) level two, application; and (c) level three, analysis, synthesis, and evaluation. In addition to these three levels, Rector totaled the criterion test scores for each level to get a score which represented the student's performance over a broad range of educational goals and cognitive behaviors.

A second factor in Rector's study was the inclusion of a classification factor dealing with the levels of mathematical ability present in the students who served as subjects. The students were divided into high and low mathematical aptitude groups on the basis of their performances on the mathematics subtest of the Scholastic Aptitude Test.

If H is used to represent those classified as having a high mathematical aptitude and a L is used to represent those having a low mathematical aptitude, the design of the study could be represented as shown in Figure 3.

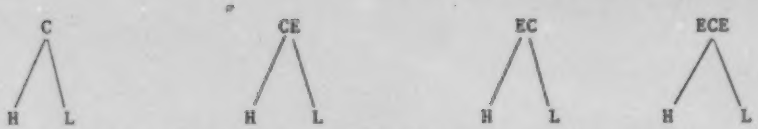


Figure 3. Design of Rector's 4 x 2 factorial research model.

The instruction in the study was carried out through the medium of programmed instruction. This mode was chosen because it allowed for complete control of the moves being used, and it equalized the personality and competency questions sometimes referred to in studies comparing teaching techniques.

The experimental materials were given to a group of undergraduate students who served as a sample, and they were allowed to work through the materials in class time. The criterion test was administered three days after the students had completed the programmed materials.

A two-way analysis of variance was run on the data from the 4 x 2 design. The analysis, run for each of the three cognitive levels and the total test scores, resulted in only one significant difference (at the .05 level), other than aptitude differences. This finding suggested that the C-strategy was significantly better than any of the other strategies in promoting student acquisition of the concepts at the Level I, knowledge and comprehension, stage of cognitive behavior. No other differences were found between the C, CE, EC, and ECE strategies or interactions between them and the levels of the aptitude factor. (These results are reported in Rector and Henderson (1970).)

Rector (1968) hypothesized that the significance of the C-strategy at the knowledge and comprehension level was due to the fact that it allowed the students to focus on the relevant material quicker and that it made fewer demands on the student than the other strategies. The other strategies, EC, CE, and ECE, required the student to infer some facts about the concept from the exemplification moves contained within the strategy. The knowledge and comprehension level of evaluation also calls for a more direct form of remembering and dealing with the information than the other levels of the taxonomy of cognitive behaviors. The lesser demands of the C-strategy on the student combined with the lesser demands of the level of evaluation of the C-strategy may be the factors which lead to the significance of the C-strategy at the Level I stage.

Rollins' Study

A second study conducted at the University of Illinois in the same general time period was the work of Rollins (1966). His study tested the relative efficacy of three strategies for teaching mathematical concepts and generalizations by guided discovery. (Rollins' study is reported in Rollins and Henderson (1967).)

Rollins proposed that a general guided discovery pattern consisted of the following decisions on the teacher's part:

1. The teacher selects a generalization with which he hopes to guide the students into discovery.

2. He selects instances of the generalization and presents them to the pupils.

3. He directs the pupils' thinking relative to the instances by means of prescriptions or leading questions.

4. He seeks evidence of abstraction by the pupils, that is seeing the common form (pattern) amid the differences. Valid evidence is quickness in giving the correct response or a statement of the correct generalization.

5. If such evidence is abundant, the teacher concludes that the pupils have discovered the generalization. If not, he repeats steps 2, 3, and 4, using new instances and perhaps employing various pedagogical aids to make the pattern more evident. (Rollins & Henderson, 1967, pp. 583-584)

Within this set of steps, various strategies for concept teaching might be identified if one views concept teaching from the standpoint of getting students to obtain the necessary and sufficient conditions for the use of the concept naming term. These different strategies would result from the selection of examples and nonexamples for steps 2 and 3 above, as well as the sequencing of these exemplification moves. The three strategies chosen for use in the Rollins study were adaptations of three inductive strategies which he ascribed to John Stuart Mill (1872). These three strategies were:

1. The stratagem of agreement. This strategy used the approach of presenting a sequence of examples from which one could infer that "Every case of p is also a case of q." Other factors in the examples would be present in some cases and absent in others, thus allowing the subjects to decide which factors are necessary or sufficient conditions for an object to be a member of the concept's referent set. Only factors p and q would be present in all cases. The students could then infer the stated generalization.

2. The stratagem of difference. This strategy used the approach of presenting a sequence of examples and nonexamples of the concept. The factors involved in these cases should lead the student to infer that "Every

case of p is also a case of q." One example might show that both p and q hold in the face of other factors, while another example shows that when the irrelevant characteristics or the prior example hold and q is absent, p is also absent. This strategy uses the logic of the contrapositive.

3. The joint stratagem of agreement and difference. This strategy is a combination of the two previous strategies and provides the student with seemingly better information on which he can infer the statement "Every case of p is also a case of q." This strategy, due to its construction from the two previous strategies, is best analyzed in the form of two different substrategies.

a. The paired instances stratagem of agreement and difference. In this substrategy, each example of the necessity and sufficient conditions obtained by the stratagem of agreement is followed by an example generated by the stratagem of difference.

b. The nonpaired instances stratagem of agreement and difference. In this substrategy, all the examples derived from one of the stratagems are presented and then all of the examples derived by the other are presented.

Rollins selected ten concepts from the secondary school geometry curriculum, which were unfamiliar to a group of junior high school students, for inclusion in a programmed instruction unit. The materials were prepared according to the stratagems of agreement, the paired instances stratagem of agreement and difference, and the nonpaired instances stratagem of agreement and difference. The three forms of concept teaching materials were designed so that the evaluation of student achievement was contained within the materials. An unreported number of instructional frames were used first then four frames were used to test the subject's attainment of the concept of concern.

These programs were then randomly assigned to a sample of eighth-grade mathematics students who had been grouped into high, average, and low ability groups on the basis of their performances on the California Short Form Test of Mental Maturity so that the three strategies were crossed with the ability groups.

The results of the evaluation frames when subjected to a two-way analysis of variance showed the following findings:

1. Students were capable of learning the concept from any of the strategies.

2. Students of high ability learned the most and students of low ability learned the least.

3. None of the hypotheses suggesting that there were no differences in the relative efficacy of the three strategies were rejected.

4. None of the hypotheses suggesting that the interactions were the same between the strategies and the ability levels were rejected.

As a result of these findings, Rollins and Henderson (1967) concluded that:

> In light of these findings it would seem that teachers, textbook writers, and programmers of automated teaching devices who wish to use inductive stratagems in teaching concepts and generalizations need not limit themselves to any one of the three stratagems investigated in this experiment. It would appear that, whichever stratagem is chosen, students of all abilities will learn from it. (p. 588)

Laboratory Studies of Concept Learning in Humans

The Search for New Domains for Relative Efficacy Studies

The foregoing studies found no significant differences of note in the relative efficacy of various patterns of moves in teaching mathematical concepts. Hence, one might wonder about the viability of the rational model of teaching for suggesting researchable hypotheses. In an effort to find situations where strategies might be especially effective, researchers examined findings of laboratory studies of human concept learning to identify typical learning problems. When such problems were identified, the laboratory findings were translated into factors in the language of the strategies of a teaching model. These factors could then be evaluated for use in relative efficacy studies from the standpoint of their logical structure.

Rosenshine and Furst (1971), in reviewing research on teacher performance criteria, suggested that more notice should be taken of the results of laboratory studies of meaningful human learning. In addition, they suggested that more attempts should be made to follow the strategy model for conducting research on teaching. Both of these suggestions seemed to indicate that more relative efficacy studies should be attempted, but with more considerations of laboratory findings and their implications for developing effective teaching strategies.

An analysis of the foregoing teaching strategy studies showed that they dealt only with the teaching of concepts in general. There was no particular effort by Rector or Rollins to identify the type of concepts being taught or whether types of teaching strategies might be differentially effective for various types of concepts. Looking at Ginther's study suggests that there might be some difference in the ways in which algebraic and geometric concepts should be handled.

Little analysis has been given to the logical types of concepts identified by Henderson in his functional taxonomy of concepts or other models of concept types developed by others.

Bruner's Analysis of Concept Types

Bruner, Goodnow, and Austin (1956) partitioned the class of concepts into three subsets: conjunctive concepts, disjunctive concepts, and relational concepts. These three types of concepts are determined by the manner in which their defining conditions are combined.

A conjunctive concept is one determined by the joint occurrence of the appropriate values of its defining attributes. It is characterized by the use of the logical connective "and." A disjunctive concept is one which is noted by the occurrence of at least one of the appropriate values of its defining attributes. It is set off by the use of the logical connective "or." The third type of concept, the relational concept, is one which is determined by an explicit relationship between the values of the defining attributes.

Henderson's Analysis of Concept Types

Henderson (1970) proposed a different partition of the class of concepts based on logical uses and properties of concepts. The general classes he offered were: denotative, nondenotative, and attributive. A denotative concept is a concept which has a nonempty referent set. A nondenotative concept, on the other hand, is a concept which has an empty referent set. The class of denotative concepts can be further partitioned into the subclasses of singular and general denotative concepts. A singular denotative concept is one with a single element in its referent set. A general denotative concept is one which has more than one element in its referent set. An attributive concept is a concept which does not have a referent set associated with it. Rather, it refers to a property that characterizes a particular set of objects, ideas, or actions. For example, one might have the idea, or concept, of "rigor" in mathematics, but the term "rigor" does not act as a sorter to partition some domain into examples and nonexamples of rigor.

Henderson (1970) further developed his functional taxonomy by differentiating between concrete and abstract concepts. A concrete concept is a denotative concept which has concrete elements in its referent set. An abstract concept is a denotative concept which does not have concrete elements in its referent set. Most of the concepts in mathematics are abstract denotative concepts. In the remainder of this paper, it is this class and its subclasses which will be the concept of concern.

In addition, this functional taxonomy of concepts can further be divided into conjunctive and disjunctive subclasses in accordance with

Bruner's classification. The results of such classifications of concepts leads to a taxonomy of concepts shown in Figure 4.

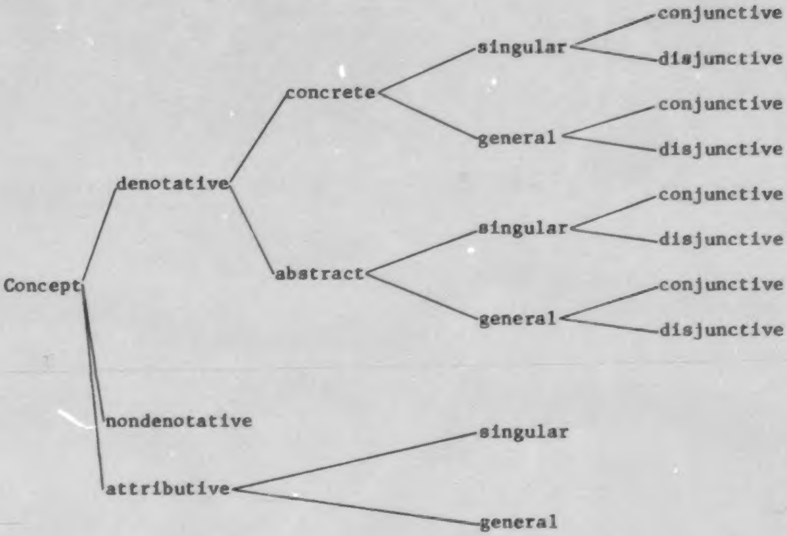

Concept
denotative
concrete
singular
conjunctive
disjunctive
general
conjunctive
disjunctive
abstract
singular
conjunctive
disjunctive
general
conjunctive
disjunctive
nondenotative
attributive
singular
general

Figure 4. The combination of Henderson and Bruner's taxonomies of concept types.

Research on Disjunctive Concept Learning

The majority of studies dealing with the nature of learners' concept attainment strategies for conjunctive and disjunctive concepts indicated that subjects encountered more difficulty with disjunctive concepts. Bruner et al. (1956) attributed the difficulty to the following reasons:

1. the inability of subjects to profit from the information contained in nonexamples of the disjunctive concepts,
2. the general tendency to avoid disjunctive concepts and disjunctive situations in life, and
3. the failure of conjunctive attainment strategies in disjunctive concept situations.

Hunt and Hovland (1960) suggested that many subjects confuse the meanings of the terms "and" and "or." In separate studies, Wells (1963) and Shore (1964) obtained results that indicated disjunctive concept attainment could be improved with training in identifying disjunctive situations. Snow and Rabinovitch (1969) compared the relative difficulty of disjunctive concepts with 9- to 13-year-olds. Their results showed that the disjunctive concepts were harder to attain at each age level than the conjunctive concepts. Snow and Rabinovitch's findings, when combined with Bruner's findings, imply that disjunctive concepts are harder to attain at all age levels.

Recent Studies of the Relative Effectiveness of Various Concept Teaching Strategies

Dossey's Studies

With these laboratory findings in mind, a study (Dossey, 1971) was designed to test the relative efficacy of four instructional strategies for teaching disjunctive concepts in mathematics. The CE, EC, and ECE strategies identified by Ginther and used by Rector were chosen as well as a CEC strategy identified in an analysis of audiotapes of classroom teaching of disjunctive concepts. The C strategy, identified and used in earlier studies, was eliminated from this study as it contained no exemplification moves. A major factor in this study was the role exemplification moves play in the subject's attainment of specified disjunctive concepts. (Dossey's study is reported in Dossey and Henderson (1974).)

The four strategies used in the study were the following:

1. Characterization-Exemplification Strategy (CE). This instructional strategy consists of four characterization moves followed by a set of six exemplification moves.

2. Characterization-Exemplification-Characterization Strategy (CEC). This instructional strategy consists of two characterization moves followed by six exemplification moves followed by an additional two characterization moves.

3. Exemplification-Characterization-Exemplification Strategy (ECE). This instructional strategy consists of three exemplification moves followed by four characterization moves followed by an additional three exemplification moves.

4. Exemplification-Characterization Strategy (EC). This instructional strategy consists of six exemplification moves followed by four characterization moves.

An example of a disjunctive concept, due to the nature of the logical structure of the concept, is an object or idea which satisfies at least one of the defining conditions of the concept. It is *not* necessary that it fulfill *all* of the disjuncts, or defining conditions, as in the case of a conjunctive concept. To be a member of a disjunctive concept's referent set it is sufficient for the object to satisfy only one of the several defining conditions for the concept. Hence, the attainment of a disjunctive concept through a sequence of positive exemplification moves amounts to a difficult information processing problem.

A nonexample of a disjunctive concept *must* interact with each of the disjuncts in the defining statement for the concept. This follows from the application of DeMorgan's law for the formation of the negative of a disjunction (Exner & Rosskopf, 1959). The resulting statement is a conjunctive statement each of whose conjuncts is a negation of one of the disjuncts of the original defining statement for the concept. This nonexample defining statement *must interact with each of the defining conditions of the disjunctive concept*. Hence, the use of nonexamples in teaching disjunctive concepts involves the relevant properties more often than the use of examples would.

In theory, it seems as if a teaching strategy for disjunctive concepts employing a large number of nonexample moves might be more efficacious than one employing a smaller number of nonexample moves. This thought resulted in the insertion of an exemplification approach factor into the study. Its levels were:

1. *Nonexample Approach* (~E). In this level of the exemplification approach factor, the ratio of nonexample to example moves is 2:1.

2. *Example Approach* (E). In this level of the exemplification approach factor, the ratio of nonexample moves to example moves is 1:2.

The insertion of the exemplification approach factor into the study suggested that nonexample moves might interact differently with students of differing intellectual ability levels; hence, a third factor, intellectual ability was added to the study. The subjects were divided into high and low ability groups with respect to the intellectual ability factor on the basis of their performance on the *Henmon Nelson Tests of Mental Maturity*. The resulting experimental design was a 4 x 2 x 2 completely crossed factorial design with 16 cells. A pictorial model of this design is given in Figure 5.

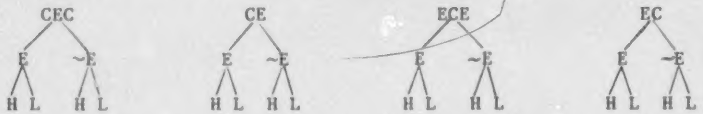


Figure 5. Pictorial model of design used in Dossey's study.

Two other points of interest were built into the study although they were not afforded a full factor status. The twelve disjunctive concepts were divided equally along the lines of algebraic and geometric concepts, as well as along the lines of whether the use of the term "or" in the concepts was in the inclusive or exclusive sense.

Programmed instructional materials were then prepared to teach the twelve contrived disjunctive mathematical concepts to the undergraduate students who served as subjects in the study. A sample page from the CEC example approach program, teaching the concept of a preve follows (Dossey, 1971).

I. A natural number which is either even or prime is called a preve. Hence, the set of all preves could be thought of as the ______________________ of the set of all even natural numbers and the set of all prime numbers.

II. 7 is a preve because it is ____________________. On the other hand, 9 is not a preve because it is not ________________ and not ___________________.

III. Give an example of a preve, other than 7, which is less than 20.____________ 15 is not a preve because it is not ____________ and it is not _________________.

IV. Tell why 47 is a preve.______________________________________
____
____

V. Give an example of a preve between 30 and 40 and tell why you believe it is a preve.____________________________________
____

VI. Explain why the set of all preves is not equal to the set of even natural numbers.______________________________________
____

VII. One could summarize the requirements for a natural number to be a preve by stating that a number must satisy what conditions?
____
____

A criterion test was constructed to measure the subjects' attainment of the concepts at each level of the cognitive behavior model developed by Rector (1968). A pilot study of the programs and tests showed that there were no significant difference in the times by the subjects to complete the different forms of the programmed units. Reliability checks for the total and cognitive level subtests showed they were appropriate for use in the experiment.

The data from the tests was submitted to a three-way analysis of variance. The analysis showed that the high ability students scored significantly higher than the low ability students at every level of comparison.

At the Level I (knowledge and comprehension) stage, significant differences existed in the exemplification approach factor. A review of the means showed that the example approach factor was more effective in promoting student attainment of concepts.

When the results of the Level II (application) stage, were studied, a significant difference was found to exist among the different strategy approaches. Duncan's New Multiple Range Test (Edwards, 1968) suggested that the CEC strategy was more effective than the ECE strategy, and all other pairs of strategy means had no significant differences between them.

At Level III (analysis, synthesis, and evaluation) stage, both the strategy and exemplification approach factors had significant differences among their component parts. The exemplification approach using four example moves and two nonexample moves was significantly more effective than the nonexample approach. On the strategy factor, the CEC, CE, and EC approaches were significantly more effective than the ECE approach. No other pairs of means for the strategy types had significant differences between them.

An analysis of the subjects' performance on the total test was carried out to examine the influence of the strategies and exemplification approaches over a wide range of educational objectives and cognitive levels. The application of Duncan's New Multiple Range Test to the strategy means indicated that the CEC and EC strategies were more effective in promoting the student learning of the disjunctive concepts than the ECE strategy. An analysis of the means for the exemplification approach factors indicated that the example approach was more effective than the nonexample approach.

An analysis of the scores on the algebraic and geometric items showed that the geometric items were significantly easier for the students to handle than were the algebraic items. The comparison of the exclusive concept scores with the inclusive concept scores showed that the exclusive concepts were significantly easier for the students to handle than were the inclusive items.

The analysis of the results on the algebraic items indicated that there was a difference in the relative efficacy of the four strategies for helping students attain the concepts. Scheffé's test for Multiple Comparisons (Winer, 1962) was then applied to the strategy treatment totals for the algebraic items. The results of several comparisons turned up only one comparison which was significant at the 0.05 level. The difference showed that the CEC strategy was significantly better than the ECE strategy in helping students attain the algebraic disjunctive concepts. No other differences were noted among the strategy treatment totals.

The analysis of the items related to the disjunctive geometric concepts indicated that a significant three-way interaction existed between the strategy, exemplification approach, and intellectual ability factors. No meaningful interpretation was made for this interaction in terms of teaching. No other significant differences were noted for the geometric items and the factors under study.

The analysis of the data corresponding to the students responses on the exclusive disjunctive items also resulted in a significant three-way interaction. Again, no meaningful interpretation was made for use in pedagogical theory. Other than this significant interaction, no other differences were judged significant in this analysis of the data from the exclusive items.

The final analysis of variance conducted on the data considered the student responses to the items involving the inclusive disjunctive concepts. Here both the strategy and exemplification approach factors had significant differences among their respective components. Scheffé's test for Multiple Comparisons indicated that the CEC strategy was more effective than the ECE strategy. In addition, the test also indicated that as a group, the CEC and EC strategies were more effective than the CE and ECE strategies as a group. The example approach of the exemplification approach factor was significantly more effective in promoting student attainment of the inclusive concepts than the nonexample approach.

The results of this study suggest several points of interest. Unlike Rector's study, several statistical differences were noted between the various strategy types and the exemplification approaches. These differences become more evident as one moves up through the levels of the cognitive behavior taxonomy. The concepts Rector used were all conjunctive in form. Hence, the differences noted might be attributed to the fact the concepts in this study were disjunctive in nature. If so, the idea that the logic of the concept is the important factor to consider should be followed up by further research.

Another factor which may have led to the differences was the change in the length of the strategies employed in the programmed materials. In Rector's study the strategies contained but five moves, while the strategies in Dossey's study contained ten moves. Does the shifting of the length of the strategy cause differences in the relative efficacy of the strategies?

Another factor which may have caused a difference was the inclusion of the CEC strategy in Dossey's study. This strategy was not used in Rector's work, and the majority of the differences in Dossey's study were related to this strategy. The common strategies for both studies showed little differences at the various levels of analyses.

The strategy differences were significant at the Level II, Level III, and Total Test analyses. The nature of the CEC strategy and its relationships to the nature of a disjunctive concept may be the cause of the superiority of this strategy. It may be the case that the early introduction of the identification move in the opening set of characterization moves in the CEC strategy fixes the concept for the student. The following set of exemplification moves allows the student to focus on the manner in which the disjuncts appear in the various examples. The last two characterization moves allow the student to focus on the relevant characteristics and properties that set off the particular disjunctive concept being studied.

The ECE format puts too much of a load on the student to initially infer the relevant defining attributes of the concept at the outset. In closing, it does not provide enough examples (three) for the student to make full use of the information in the characterization moves in discriminating between examples and nonexamples of the concept. It also does not provide the same type of closure the CEC strategy does. This closure of characterization moves may be more important than the composition of the set of initial moves in the strategy. In the analysis of the relative efficacy of the strategies for dealing with inclusive concepts, the CEC and EC strategies, those closing with sets of characterization moves were judged to be more effective as a group than those ending in exemplification moves, namely CE and ECE.

The same statements apply to the CE strategy that applied to the ECE strategy. The EC strategy seemed to function fairly well in that it was always ranked second to the CEC strategy in terms of effectiveness. Its performance might be attributed to the larger span of time the student had to infer, from the sequence of exemplification moves, the nature of the disjunctive concept. In addition, it had the closure section of a group of characterization moves. The fact that the order of the means for the four strategies at all levels of analysis was the same, namely ECE, CE, EC, CEC, moving from low to high suggests the import of closure with the characterization moves.

The exemplification approach differences at both the Level I and Total Test showed that the predominance of example moves in a strategy was more efficacious than a predominance of nonexample moves. One might theorize that the example moves were more important for the cognitive requirements of the items on the Level I exam than they were for the other levels of evaluation. The differences might also be explained by agreeing with Bruner's claim that students have more difficulty in dealing with nonexamples of a concept.

The significant role played by the strategies in dealing with algebraic concepts, while no significant differences were found in the relative efficacy of the strategy factor for the geometric concepts, might be due to the lack of visual elements in the moves for algebraic concepts. Example moves for geometric concepts usually involve a visual image while the algebraic moves may or may not involve such a representation. Further research needs to be done to clarify this issue. Some combination of the present findings and Ginther's work might serve as a starting point.

The findings on the comparison of the exclusive and inclusive test item scores showed differences on both the strategy and exemplification factors for the inclusive concept items. The use of the example moves may have been more effective here because these items can satisfy one or both of the defining conditions at once. The strategy factor differences might be attributed to the closure moves of the CEC and EC strategies. This closure set of moves may help the student overcome the confusion of the "and" and "or" terms in relation to the nature of inclusive disjunctive concepts.

Malo's Study

Two additional relative efficacy studies have been carried out by educational researchers (Gaston & Kolb, 1973; Malo, 1974). We shall examine the design and results of Malo's study first as it relates to both the Rollins and Dossey studies.

Malo compared the relative efficacy of five exemplification strategies to teach the twelve contrived disjunctive concepts used in Dossey's study to a group of undergraduate students. In addition, he prepared a programmed instruction unit which was to teach students how to make use of the information contained in exemplification moves in the attainment of concepts. Some emphasis was placed on the use of nonexamples in this process.

Malo's study made use of the levels of cognitive behavior developed by Rector to analyze the relative efficacy of the various strategies at different levels of intellectual functioning. He also administered the Henmon Nelson Tests of Mental Maturity to obtain a covariate measure to use in the analysis of the relative efficacy of the strategies.

Results of studies in psychological concept attainment (Hunt, 1962) and the results of the exemplification approach factor in Dossey's (1971) study suggested to Malo the following five exemplification strategies (Malo, 1974, p. 26):

1. Alternating Example and Nonexample Strategy (EN). This strategy consists of starting with an example and successively alternating six examples with six nonexamples.

2. Alternating Nonexample and Example Strategy (NE). This strategy consists of starting with a nonexample and successively alternating six nonexamples with six examples.

3. Alternating Grouped Examples and Nonexamples Strategy (6EN). This strategy consists of six examples followed by six nonexamples.

4. Alternating Grouped Nonexamples and Examples Strategy (6NE). This strategy consists of six nonexamples followed by six examples.

5. Grouped Examples Strategy (E). This strategy consists of twelve examples.

Programmed instructional units were then developed to teach the twelve concepts to the students in the sample. The materials were carefully developed to keep the information contained in the moves balanced with respect to information about the disjuncts in the defining statements for the concepts. The materials were then given to students who were allowed to complete them in class time.

The design for Malo's study is shown in Figure 6. The students in the groups marked with a "∼U" were the students who did not study the unit on learning disjunctive concepts from exemplification moves. The students in the groups marked with a "U" were the ones that used the experimental units.

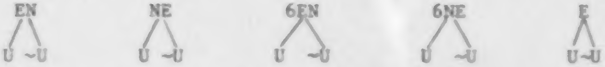


Figure 6. The experimental design used for Malo's study.

The analyses performed on the total test data showed that there were no significant differences between the five different exemplification strategies in promoting student attainment of the disjunctive mathematical concepts. Malo conjectured that the absence of the differences might be a result of the length of the strategies. Another possibility cited was the usage of programmed instruction. It might be possible that the use of this instructional technique might wipe out some differences that might hold with a teacher's use of the strategies.

The analysis of the student responses for the different levels of cognitive behavior taxonomy revealed only one significant difference among the strategies. At Level II (application) the Grouped Examples Strategy was shown to be significantly better than the remaining strategies. This finding was consistent with Bruner's findings (Bruner et al., 1956). When the adjusted mean scores for the other strategies were examined, strategies that alternated examples with nonexamples seemed to have some advantage over those that grouped both the examples and the nonexamples.

The analysis of the effects of the experimental unit on learning from exemplification moves showed that no significant difference could be attached to the use of the unit. The lack of a difference here may be due to several factors. One is that the students already knew how to make use of such information, while another is that the strategy lengths might have destroyed any differences which resulted from the use or nonuse of the unit. A third explanation might be that the concept teaching strategies were so carefully constructed that they taught the students in both groups the techniques for learning from examples. If this was the case, the differences due to the use of the experimental units would have been wiped out.

Two other findings of Malo's study were similar to results noted in Dossey's study. Exclusive disjunctive concepts were significantly easier for students to attain than inclusive disjunctive concepts. In addition, the geometric concepts were significantly easier for students to handle than were the algebraic disjunctive concepts.

Gaston and Kolb's Study

Gaston and Kolb (1973) compared the relative effectiveness of three strategies for teaching the concept of a partition of a set. The study employed programmed instructional materials to instruct a sample of undergraduate students about the concept. Bidwell (1974) has commented that this study had some methodological problems; however, the design and factors chosen offer some suggestions for further research. The following three strategies were used in this study:

1. The first strategy was a four move CE strategy consisting of an identification move followed by two example moves with justification and then a third example move with instructions for the student to verify that the example truly was an example.

2. The second strategy was an ECE strategy which opened with three example moves. These moves were followed with six single characteristic moves which focused on the relevant defining conditions for a set partition. After the characteristic moves, the student was given the original three exemplification moves again and then four more new exemplification moves, three of which were examples and one of which was a nonexample. This resulted in a strategy sixteen moves in length.

3. The third strategy was a thirty-eight move exemplification strategy consisting of moves which required the student to infer what a partition was and then discriminate between partitions and other subdivisions of a set.

To evaluate the relative effectiveness of the three different strategies on student learning, three tests were given. The first was a Vertical Transfer Test (Gagné, 1970). This test was aimed at measuring the students' ability to transfer their knowledge about partitions from the examples and characteristics they had seen to ranking a list of generalizations concerning partitions as being sometimes, always, or never true. The ten item Vertical Transfer Test was followed by a thirty item test called the Exemplification Test. This second test required the subject to judge whether or not a particular set subdivision was or was not an example of a partition of a set. The third test, called the Characterization Test, consisted of fifteen statements which supposedly characterized the concept of a partition of a set. The student was then required to label the statement given as being true or false.

An analysis of the results of the study, while hampered by the small numbers and the inability to combine the scores from the two classes used, suggested that there was no significant difference between the effectiveness of the strategies on either the Vertical Transfer or the Characterization Test. However, the results on the Exemplification Test when analyzed by orthogonal contrasts indicated that the mean of the E-strategy group was significantly greater than the pooled means of the CE and ECE groups. No significant difference existed between the means of the CE and ECE groups.

The concept of a set partition is a conjunctive concept. Hence, the results from this study can be compared and contrasted to the results identified by Rector and by Rollins. Rector's findings suggested that the only differences occurred at the knowledge and comprehension level of evaluation. His results showed that the C-strategy was most effective at this level. Gaston and Kolb's study did not have any test that measured cognitive activity at this level, and they did not have a C-strategy. Gaston and Kolb found that an E-strategy was most effective on an Exemplification Test. This finding would compare with a finding at Rector's Level II, but Rector did not have an E-strategy. Neither study showed significant differences existing at the higher levels of cognitive behavior.

In considering these studies, one is led to the conjecture that there may be some training effect between the type of strategy used and the type of evaluation used. Support for this conjecture is given by Rector's C-strategy showing up well on the Level I test and Gaston and Kolb's E-strategy winning out on the Exemplification Test.

Directions for Future Relative Efficacy Studies in Mathematical Concept Teaching

A Review of Known Results

Comparisons among the foregoing studies may not be too accurate in that there were many differences in the models used. A crucial difference was in the lengths of the strategies employed. Rector's strategies were five moves in length, Dossey's were ten moves long, Malo's were twelve moves long, and Gaston and Kolb's ranged from four to thirty-eight moves in length. Future studies should be careful to note the effects of different strategy lengths and the amount of time the subjects devote to the learning process, as these factors may have a significant effect on the efficacy of a particular strategy. Some efforts should be made to keep the evaluation designs, or some aspect of them, similar enough to permit some form of comparison between studies. Such studies might also allow for further analyses of the questions considered before while offering partial replications of the prior studies.

Future studies should attempt to take the following results of the previously mentioned studies into account in order that they may be considered again in both similar and different settings:

1. Different logical forms of concepts may affect the relative efficacy of concept teaching strategies (Dossey, 1976).

2. Exemplification strategies do not differ significantly among themselves, but may be quite effective in preparing students to function with conjunctive concepts at the application level (Gaston & Kolb, 1973; Malo, 1974; Rollins, 1966).

3. Characterization strategies seem to be very effective in promoting student achievement of conjunctive concepts at the knowledge and comprehension level (Rector, 1968).

4. Differences exist in students' ability to handle algebraic and geometric disjunctive concepts, as well as in their ability to deal with inclusive and exclusive disjunctive concepts (Dossey, 1971; Malo, 1974).

5. Strategies and exemplification approaches differ in their abilities to handle algebraic and inclusive disjunctive concepts (Dossey, 1976).

6. High ability students did significantly better than low ability students no matter what concept teaching strategy was employed.

These findings do answer several questions concerning the conditions under which various concept teaching strategies are effective. However, they still do not form a firm foundation for making pedagogical decisions concerning concept teaching.

Shortcomings of the Relative Efficacy Studies Reviewed

Several criticisms have been made concerning the manner in which the foregoing studies have been conducted. Swank (1973) suggests that the use of programmed instruction limits the generalizability of the results of the studies to any form of classroom teaching. He also mentions that the strategies employed were too short, i.e., they contained too few moves. Sowder (1974) questioned the use of contrived concepts, such as were used in both Dossey's and Malo's studies. A later study (Sowder, 1975) indicates that there may be no real problems in projecting findings from studies using contrived concepts.

The Use of the Findings of Relative Efficacy Studies

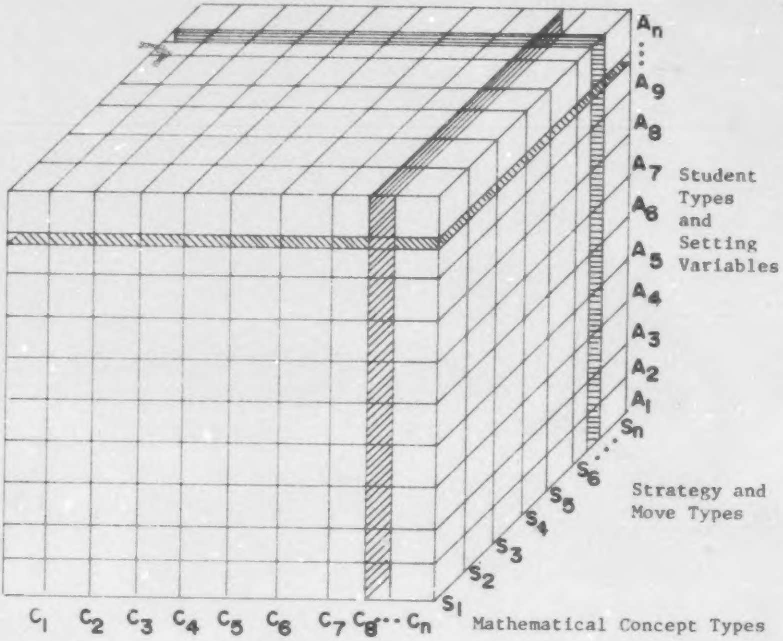
With these limitations, the findings of these tightly controlled studies do provide some directions for the mathematics educator. The results suggest that some strategies may be more effective for developing certain types of concepts in writing textual materials, programmed instruction units, or computer assisted instruction materials. In addition, they provide some direction for determining the conditions under which classroom studies of the relative efficacy of various concept teaching strategies might find significant differences. Carefully designed studies of classroom teaching using different strategies might be carried out to attempt to replicate the studies which have been carried out via programmed instruction. Swank (1973) and Benjamin (1971) have carried out two studies comparing concept teaching strategies in regular classroom settings.

In addition, the results of the controlled analyses of concept teaching strategies might be used in microteaching situations or in the development of protocol materials on concept teaching. Such studies might also provide empirical findings which can be used to justify the study of concept teaching strategies in mathematics education texts, for example, Dynamics of Teaching Secondary School Mathematics (Cooney et al., 1975). Further, the findings may stimulate further studies on the relative efficacy of concept teaching strategies.

A Model for Further Research

The design of future studies might consider the following research paradigm for concept teaching strategies. This model is developed from the model proposed by Henderson (1970) and Turner (in this monograph).

The model, pictured in Figure 7, draws its major dimensions from Turner's suggestions. The first dimension consists of the variety of concepts considered in the mathematics curriculum from grades K-12. The various concepts might be listed individually or they might be divided into various subdivisions according to various classificatory rules such as: general-vague, denotative-nondenotative, algebraic-geometric, conjunctive-disjunctive, or singular-general.



Aₙ
A₉
A₈
A₇
A₆
A₅
A₄
A₃
A₂
A₁
Student Types and Setting Variables
Sₙ
S₆
S₅
S₄
S₃
S₂
S₁
Strategy and Move Types
C₁ C₂ C₃ C₄ C₅ C₆ C₇ C₈ ... Cₙ
Mathematical Concept Types

Figure 7. The model for future concept teaching research.

The second dimension of the model, teacher actions, is concerned with various factors believed to be relevant to student acquisition of the concepts under study. Turner suggests subdivisions along this dimension to consist of both strategies and moves. A researcher must consider the impact of various strategy types on student learning, as well as the relative power of individual types of moves. The third dimension is the one that considers the various attributes possessed by students. These attributes are believed to be correlated with students' ability to attain the conceptual material of interest. Such factors might be achievement, aptitude, or attitude scores from tests or other sources of information. A fourth dimension which is not pictured in Figure 7 consists of various levels of indicators of student attainment of the concepts of interest. Such indicators might take the form of a taxonomy of levels of mathematical cognitive achievement.

Topics for Further Investigation

Some topics which need further investigation are:

1. Contemporary texts could be sampled at the elementary and secondary levels to determine what strategies are being used. Such a study might go along the directions laid out by Ginther. An interesting point would be analysis of the algebra and geometry concept teaching strategies employed by the same author in the cases where texts by the same author exist.

2. The "power" of a single move could be examined with respect to another move at the same point. For example, does an identification move and a necessary condition move have the same effect on student acquisition of a concept. Here strategies identical except for the one move would have to be used.

3. Studies investigating the role of telling and the role of questioning also fall into the realm of concept teaching strategies in mathematics.

4. The length of strategies and their relationship to student attainment of concepts is also another area of interest (Dossey, 1975). If the design of this type of research calls for evaluation to take place within the strategy, the investigator must be careful to realize he is adding a move to the strategy at this point.

5. More work needs to be done in examining the role of example and nonexample moves in concept teaching. Shumway (1974) has done work that provides insight in this area. Such studies must carefully balance the total amount of information conveyed in the strategies.

The findings from such studies would, when combined with extant results, move us toward a theory of mathematical concept teaching which would begin to take on the characteristics suggested by Gage (1963) and Bruner (1966). They would also provide the methods teacher with a set of generalizations concerning the teaching of mathematical concepts drawn from empirical studies of the relative efficacy of concept teaching strategies.

References

Bidwell, J. K. Review of A comparison of three strategies for teaching a selected mathematical concept to students in college algebra by J. A. Gaston & J. R. Kolb. Investigations in Mathematics Education, 1974, 7, 25-27.

Bloom, B. S., Englehart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. Taxonomy of educational objectives handbook I: Cognitive domain. New York: David McKay, 1956.

Benjamin, M. P. Comparative study of two strategies for teaching measurement concepts. Unpublished master's thesis, Illinois State University, 1971.

Bruner, J. S. Toward a theory of instruction. New York: Norton, 1966.

Bruner, J. S., Goodnow, J. J., & Austin, G. A. A study of thinking. New York: Wiley, 1956.

Carroll, J. B. Words, meaning and concepts. Harvard Educational Review, 1964, 34, 178-202.

Cooney, T., Davis, E., & Henderson, K. B. Dynamics of teaching secondary school mathematics. Boston: Houghton Mifflin, 1975.

Dossey, J. A. The relative effectiveness of four strategies for teaching disjunctive concepts in mathematics (Doctoral dissertation, University of Illinois, 1971). Dissertation Abstracts International, 1972, 32A 5467-5468. (University Microfilms No. 72-12, 142)

Dossey, J. A. The relative effectiveness of four strategies for teaching algebraic and geometric disjunctive concepts and for teaching inclusive and exclusive disjunctive concepts. Journal for Research in Mathematics Education, , 1976, 7, 92-105.

Dossey, J. A. The differential effect of strategy length in concept teaching strategies for disjunctive mathematical concepts. Paper in preparation, 1975.

Dossey, J. A., & Henderson, K. B. The relative effectiveness of four strategies for teaching disjunctive concepts in mathematics. Journal for Research in Mathematics Education, 1974, 5, 6-19.

Edwards, A. L. Experimental design in psychological research (3rd ed.). New York: Holt, Rinehart & Winston, 1968.

Exner, R. M., & Rosskopf, M. F. Logic in elementary mathematics. New York: McGraw-Hill, 1959.

Gage, N. L. Paradigms for research on teaching. In N. L. Gage (Ed.), Handbook of research on teaching. Chicago: Rand McNally, 1963.

Gagné, R. M. The conditions of learning. New York: Holt, Rinehart & Winston, 1970.

Gaston, J. A., & Kolb, J. R. A comparison of three strategies for teaching a selected mathematical concept to students in college algebra. Journal for Research in Mathematics Education, 1973, 4, 177-186.

Ginther, J. L. A study of definitions in high school mathematics textbooks (Doctoral dissertation, University of Illinois, 1964). Dissertation Abstracts, 1965, 25, 4574. (University Microfilms No. 65-00818)

Ginther, J. L., & Henderson, K. B. Strategies for teaching concepts by using definitions. The Mathematics Teacher, 1966, 59, 455-457.

Henderson, K. B. A logical model for conceptualizing and other related activities. Educational Theory, 1963, 13, 277-284.

Henderson, K. B. A model for teaching mathematical concepts. The Mathematics Teacher, 1967, 60, 573-577.

Henderson, K. B. Concepts. In M. F. Rosskopf (Ed.), The teaching of secondary school mathematics. Thirty-third Yearbook of the National Council of Teachers of Mathematics. Washington, D.C.: The Council, 1970.

Henderson, K. B. The development and use of pedagogical models. Paper presented at the annual meeting of the American Educational Research Association, Chicago, April 1972.

Hunt, E. B. Concept learning: An information processing problem. New York: Wiley, 1962.

Hunt, E. B., & Hovland, C. Order of consideration of different types of concepts. Journal of Experimental Psychology, 1960, 59, 220-225.

Malo, G. E. Differential treatments in learning disjunctive concepts in mathematics (Doctoral dissertation, University of Illinois, 1974). Dissertation Abstracts International, 1975, 35A, 4051. (University Microfilms No. 75-363)

Mill, J. S. A system of logic, ratiocinative and induction, being a connected view of the principles of evidence and the methods of scientific investigation. New York: Harper, 1872.

Nuthall, G., & Snook, I. Contemporary models of teaching. In R. M. W. Travers (Ed.), The second handbook of research on teaching. Chicago: Rand McNally, 1973.

Rector, R. E. The relative effectiveness of four strategies for teaching mathematical concepts (Doctoral dissertation, University of Illinois, 1968). Dissertation Abstracts, 1968, 29A, 520. (University Microfilms No. 68-12,184)

Rector, R. E., & Henderson, K. B. The relative effectiveness of four strategies for teaching mathematical concepts. Journal for Research in Mathematics Education, 1970, 1, 69-75.

Rollins, J. H. A comparison of three stratagems for teaching mathematical concepts and generalizations by guided discovery (Doctoral dissertation, University of Illinois, 1966). Dissertation Abstracts, 1966, 27A, 711-712). (University Microfilms No. 66-7802)

Rollins, J. H., & Henderson, K. B. A comparison of three stratagems for teaching mathematical concepts and generalizations by guided discovery. The Arithmetic Teacher, 1967, 14, 583-588.

Rosenshine, B., & Furst, N. Research on teacher performance criteria. In B. O. Smith (Ed.), Research in teacher education: A symposium. Englewood Cliffs, N. J.: Prentice Hall, 1971.

Shore, E. Reasoning ability and stimulus variables in learning disjunctive concepts (Doctoral dissertation, The Pennsylvania State University, 1963). Dissertation Abstracts, 1964, 24, 3441. (University Microfilms No. 64-1416)

Shumway, R. J. Negative instances in mathematical concept acquisition: Transfer effects between the concepts of commutativity and associativity. Journal for Research in Mathematics Education, 1974, 5, 197-211.

Sowder, L. Review of The relative effectiveness of four strategies for teaching disjunctive concepts in mathematics by J. Dossey & K. B. Henderson. Investigations in Mathematics Education, 1974, 7, 17-20.

Sowder, L. Concept attainment problems in two settings. Paper presented at the annual meeting of the National Council of Teachers of Mathematics, Denver, April 1975.

Smith, B. O. On the anatomy of teaching. Journal of Teacher Education, 1956, 7, 339-346.

Smith, B. O., Meux, M., Coombs, J., Nuthall, G., & Precians, R. A study of the strategies of teaching. Urbana: Bureau of Education Research, University of Illinois, 1967.

Smith, B. O., Meux, M., Coombs, J., Eierdam, D., & Szoke, R. A study of the logic of teaching. Urbana: University of Illinois Press, 1970.

Snow, C., & Rabinovitch, S. Conjunctive and disjunctive thinking in children. *Journal of Experimental Child Psychology*, 1969, *7*, 1-9.

Swank, E. W. A comparison of selected strategies for the teaching of mathematical concepts (Doctoral dissertation, The University of Georgia, 1973). *Dissertation Abstracts International*, 1974, *34A*, 5001. (University Microfilms No. 74-4887)

Wells, H. Effects of transfer and problem structure in disjunctive concept formation. *Journal of Experimental Psychology*, 1963, *65*, 63-39.

Winer, B. J. *Statistical principles in experimental design*. New York: McGraw-Hill, 1962.

An Empirical Comparison of Teaching Strategies Where the Amount of Content Information and Teacher-Pupil Interaction Is Varied[1]

Earl W. Swank

Coastal Plains Cooperative Educational Services Agency

Valdosta, Georgia

Teacher talk is an integral, if not essential, aspect of teaching strategies (Hughes, 1963; Meux & Smith, 1964). Since teacher talk represents a sizable portion of the classroom dialogue it is justifiably the concern of research. Certainly, as the teacher makes verbal contributions to the classroom dialogue, he is communicating information about content. Gage (1972) wrote that the substantive content in the teacher-pupil interactions should have some impact on the learning that the students experience.

Research concerned with describing the substantive part of the teaching act has been carried out by Bellack (1965) and Smith, Meux, Coombs, Nuthall, and Precians (1967). A logical outgrowth of studies such as the two just mentioned is to manipulate the substantive aspect of the classroom dialogue in an experimental setting. The present study is an example of how the substantive content relative to selected mathematical concepts can be manipulated in a simulated classroom setting and how such manipulations affect student learning.

Statement of the Problem

One purpose of this study was to determine if student achievement is sensitive to variations in the amount of content information. A second purpose was to investigate the effect of teacher-pupil verbal interaction on student achievement. The amount of content information and the amount of teacher-pupil verbal interaction in establishing the content information were regulated simultaneously so that possible interactions of these two factors could be ascertained. A student ability and treatment interaction effect was made possible for investigation by randomly selecting students on two different ability levels for participation in the study.

---

[1] This paper is based on a doctoral dissertation submitted to the Department of Mathematics Education, University of Georgia, in 1973.

One variable under consideration was concerned with the amount of content information contained in the classroom discourse relative to a specific concept. The amount of content information that is transmitted by an instructional strategy can be described by the number of concept moves. Operationally, a concept move as described in Henderson's (1967) model for the teaching of mathematical concepts can be thought of as a unit of content information. Certainly, different types of concept moves provide different kinds of content information, but with careful limitations the number of concept moves can serve as an indicator of the amount of content information contained in an instructional strategy. Emphasis for studying content information contained in the classroom discourse is provided by Gage (1972) who stated, "By all that is plausible the logical and substantive content of the classroom content ought to have some connection with knowledge and comprehension students acquire" (p. 313). A second variable was the amount of teacher-pupil verbal interactions in establishing the content information within the classroom discourse. As the concept moves are established in the classroom dialogue, there is some probability that the students will be cognitively involved. However, it seems reasonable to assume that "the probability that students are cognitively involved is directly proportional to the amount of overt participation" (Snow, 1970, p. 25). Since verbal responses represent one category of overt participation, the verbal responses can be used to indicate when students are cognitively involved in the establishment of a concept move.

If the set of teacher-pupil verbal interactions is restricted to the verbal interactions related to the establishment of concept moves, then from a learning viewpoint it appears that the value of a concept move to a student is enhanced if the student contributes something to the establishment of that concept move. Therefore, the independent variable, concept move interaction, was defined to systematically control whether a concept move was a result of teacher talk or a combination of teacher talk and student talk. A classification system was developed to distinguish three different types of concept move interactions that may occur and is given in Table 1. A method of describing distinct levels of concept move interactions is provided by this classification system.

An example is provided to illustrate the difference between categories 1 and 2. If the teacher names a pairing of the members in two sets and then asks the students if this pairing is a function, some students may say "Yes, it is a function." If the teacher goes on to something else or follows the student response with a justification of why the pairing is a function, the concept move interaction would be classified as a "1." However, if the teacher had asked why it is a function and the student gave the justification, the concept move interaction would be classified as a "2."

Table 1

Concept Move Interaction Categories

Category	Amount of Concept Move Interaction
0	The teacher completes the concept move in its entirety.
1	The teacher completes the concept move with a short response from the student.
2	The teacher completes the concept move with either (1) two or more short responses or (2) a lengthy response by one or more students.

By operationally defining two levels for each of the two variables, frequency of concept moves and concept move interaction, it was possible to develop four instructional strategies. The four instructional strategies, denoted HH, HL, LH, and LL, are described below:

1. HH is an instructional strategy employing a relatively high frequency of concept moves combined with a high amount of teacher-pupil verbal interaction in establishing the concept moves.

2. HL is an instructional strategy employing a relatively high frequency of concept moves combined with a low amount of teacher-pupil verbal interaction in establishing the concept moves.

3. LH is an instructional strategy employing a relatively low frequency of concept moves combined with a high amount of teacher-pupil verbal interaction in establishing the concept moves.

4. LL is an instructional strategy employing a relatively low frequency of concept moves combined with a low amount of teacher-pupil verbal interaction in establishing the concept moves.

Instructional Strategies

The two levels of concept move interaction were defined in the following manner. A high concept move interaction strategy contained at least twice as many moves in the "2" classification as in the "0" classification. Reversing this ratio of moves in the "0" and "2" classifications defined the low concept move interaction strategy. In both strategies the number of concept move interactions classified as "1" were minimized. Tables 2A and 2B contain the number of concept moves in each of the four strategies and the distribution of concept move interactions planned for each strategy.

As the frequencies in Tables 2A and 2B indicate, the number of concept moves in the high frequency concept move strategies (HH, HL) is approximately twice the number of concept moves in the low frequency concept move strategies (LH, LL). Although there was a relatively large difference between the two frequencies of concept moves, there was no difference in the types of concept moves. To illustrate, consider one particular type of concept move, say the "example" concept move. There may be two examples of functions represented by an arrow diagram in the low frequency concept move strategy so the high frequency concept move strategy would contain four examples of functions using the arrow diagram notation. Thus, the 2 to 1 ratio was preserved across each type of concept move for each of the three concepts taught--function, inverse function, and constant function.

Table 2A

Planned Frequencies of Concept Move Interactions (High Frequency Strategy)

HH Strategy		HL Strategy	
Interaction Categories	Number of Moves (N)	Interaction Categories	Number of Moves (N)
0	$N \leq 36$	0	$N \geq 74$
1	$N \leq 11$	1	$N \leq 11$
2	$N \geq 74$	2	$N \leq 36$
Total	121	Total	121

Table 2B

Planned Frequencies of Concept Move Interactions (Low Frequency Strategy)

LH Strategy		LL Strategy	
Interaction Categories	Number of Moves (N)	Interaction Categories	Number of Moves (N)
0	$N \leq 14$	0	$N \geq 36$
1	$N \leq 11$	1	$N \leq 11$
2	$N \geq 36$	2	$N \leq 14$
Total	61	Total	61

Description of Tests

An achievement test was constructed and field-tested in a pilot study prior to the implementation of the present study. An item analysis of the achievement test led to the elimination of those items too easy or too difficult or that had negative discrimination values. Another revision in the achievement test was accomplished by including items concerned with the composition of functions even though this concept was not included in the instructional strategies. The composition items were introduced to provide some measure of the relative transfer effect of each instructional strategy.

Each item on the revised achievement test was classified by three judges according to the cognitive behavior required for a successful answer. Four cognitive levels based on Bloom's taxonomy were used for classifying each of the test items. The four cognitive levels and the frequency of items at that level were: knowledge (6), comprehension (22), application (25), and analysis (15). The 68 test items included 16 true-false, 23 multiple choice, and 29 completion-type questions. A parallel form of the achievement test was constructed by using equivalent item forms, thus one form served as the posttest and the other as the retention test. The posttest was administered one day after the completion of the instructional strategies, and the retention test was given one month later. There was no time restriction although everyone finished in less than one hour.

Reliability coefficients were calculated for the posttest, retention test, and the cognitive subtests using the Kuder-Richardson formula 20 (KR20). Table 3 contains the calculated reliability values.

As indicated in the table, the reliability values for the total tests were adequate while the knowledge and analysis subtests were less than desirable. These two subtests contained the fewest items (6 and 15 items respectively) and also the least variance.

Table 3

Reliability Coefficients

Posttest		Retention Test	
Total	.77	Total	.77
Knowledge Items	-.02	Knowledge	.79
Comprehension Items	.65	Comprehension	.65
Application Items	.66	Application Items	.58
Analysis Items	.13	Analysis Items	.19

Hypotheses

The following null hypotheses were tested in the study.

H_1: There is no difference between the mean performance of students experiencing the high frequency of concept moves and the mean performance of students experiencing the low frequency of concept moves on the posttest achievement measure.

H_2: There is no difference between the mean performance of students experiencing the high level of concept move interaction and the mean performance of students experiencing the low level of concept move interaction on the posttest achievement measure.

H_3: There is no difference between the mean performance of the ability level I students and the mean performance of the ability level II students on the posttest achievement measure.

h_4: There is no significant interaction of the frequency of concept moves and concept move interaction on the posttest achievement measure.

H_5: There is no significant interaction of concept move interactions with student abilities on the posttest achievement measure.

H_6: There is no significant interaction of frequency of concept moves with student abilities on the posttest achievement measure.

$H:_7$ There is no significant interaction of frequency of concept moves, concept move interaction, and the student ability levels on the posttest achievement measure.

Hypotheses H_8 through H_{35} were generated by replacing the word "posttest" in the seven hypotheses above with each of the following phrases:

1. the knowledge items subtest of the posttest,
2. the comprehension items subtest of the posttest,
3. the application items subtest of the posttest, and
4. the analysis items subtest of the posttest.

Hypotheses H_{36} through H_{70} were generated by using each of the retention measures listed below as the dependent variable in place of the posttest measure in hypotheses H_1 through H_7:

1. the total retention test measure,
2. the knowledge items subtest measure of the retention test,
3. the comprehension items subtest measure of the retention test,
4. the application items subtest measure of the retention test, and
5. the analysis items subtest measure of the retention test.

The null hypotheses listed above were tested to indicate which factors or combinations of factors significantly affect student achievement as measured by a posttest and retention test. The treatment interaction hypotheses were of particular interest since the combining of treatments might have an effect different than that expected from considering each treatment independently (Winer, 1962).

Sample

The subjects selected for participation were chosen from the eighth-grade class at Clarke Middle School in Athens, Georgia. Clarke Middle School is an integrated public school with approximately 330 students in the eighth-grade from all socio-economic levels.

School records were used to obtain I.Q. measures (California Test of Mental Maturity) which were then used to select students to pretest for possible participation in the study. Since average and above average students were desired for participation, it was arbitrarily determined to pretest only students with an I.Q. measure of 100 or higher.

Approximately 125 students were pretested to determine the selection of students for participation in one of the four instructional strategies. The pretest items were partitioned into three subtests corresponding to the first three stages of Thomas' hierarchy (1975)[2] which permitted the placement of each student at one of the three stages. All students placed at stage two or higher were eliminated from further consideration for participation in the study. After the eliminations were completed, the next procedure was to separate the remaining students into two distinct ability groups on the basis of their I.Q. measures. This separation process was determined by the distribution of I.Q. measures among the remaining students participating in the study.

Ability level I students were defined to be students with an I.Q. measure of 100 to 105 inclusive. Ability level II students were defined to be students with I.Q. measures of 111 to 125 inclusive. It was necessary to use a wider range of I.Q. measures for the ability level II students to obtain an adequate number of students.

The final selection procedure was to randomly select five students from the ability level I group and five from the ability level II group to form an instructional group. This randomization and stratification procedure was repeated until eight groups of ten students were selected. Each instructional strategy was then randomly assigned to two of the eight instructional groups. This procedure permitted the replication of each instructional strategy.

Analysis of Concept Moves and Concept Move Interactions

To insure fidelity of the instructional strategies implemented to the planned instructional strategies, each instructional session was audio-recorded and analyzed in terms of concept moves and concept move interactions. The analysis of the taped lessons allowed a comparison of the observed concept moves and concept move interactions with the planned concept moves and concept move interactions. Analysis of the taped lessons was done daily by the investigator to obtain daily feedback to indicate the rigor with which each strategy was implemented.

---

[2] A sequence of five stages was developed by Thomas to describe the development of the function concept. For a complete description and discussion of the stages see Thomas (1975).

A sequence of five twenty minute lessons was employed to implement each of the four instructional strategies. The investigator was the only teacher implementing the four instructional strategies. Approximately two weeks were required for implementation of the instructional strategies and the administration of the posttest.

A comparison of the distribution of planned concept moves with the distribution of observed concept moves demonstrates that the planned lessons were implemented. The distributions for the planned and observed concept moves related to the three selected concepts--functions, inverse function, and constant function--for each instructional strategy are given in Table 4. The data in Table 4 indicate the number of observed moves is equal to or less than the planned moves. Some of the planned concept moves were stated improperly or not completed which resulted in the number of observed moves being slightly less than the number of planned concept moves.

The taped lessons were also analyzed in terms of the concept move interactions to determine how well the different concept move interaction strategies were implemented. The number of observed concept move interactions and the number of planned concept move interactions for each instructional strategy are given in Table 5.

The frequencies of the observed concept move interactions were in the desired direction compared to the frequencies of the planned concept move interactions except for three instances--interaction categories 1 and 2 for the HH_1 strategy and category 1 for the HH_2 strategy. Despite the three discrepencies noted, the data indicate that two distinct levels of concept move interaction were implemented.

Data Analysis

The null hypotheses were tested using univariate analysis of variance (ANOVA) procedures. A three-way analysis of variance procedure was used on the posttest and retention test measures to test the equality of means across each of the variables--frequency of concept moves, concept move interaction, and student abilities. The three-way ANOVA also permitted the three two-factor interactions and one three-factor interaction to be tested for significance. A three-factor experimental design with each factor fixed is referred to as a Model I design (Winer, 1962, p. 172).

Table 4

Distributions of Observed and Planned Concept Moves

Concepts	Function		Inverse Function		Constant Function	
Instructional Strategy[a]	Observed	Planned	Observed	Planned	Observed	Planned
HH_1	87	90	23	29	13	12
HH_2	87	90	26	29	10	12
HL_1	86	90	26	29	11	12
HL_2	87	90	24	29	11	12
LH_1	45	45	11	13	6	7
LH_2	41	45	11	13	5	7
LL_1	45	45	11	13	6	7
LL_2	44	45	11	13	5	7

[a] Each instructional strategy was implemented twice, so HH_1 indicates the frequency of moves observed in the first implementation and HH_2 the frequency observed in the second implementation.

Table 5

Observed and Planned Concept Move Interactions

Interaction Categories	0		1		2	
Instructional Categories	Observed	Planned	Observed	Planned	Observed	Planned
HH_1	31	$N \leq 36$	20	$N \leq 11$	72	$N \geq 74$
HH_2	21	$N \leq 36$	26	$N \leq 11$	76	$N \geq 74$
HL_1	97	$N \geq 74$	15	$N \leq 11$	9	$N \leq 36$
HL_2	94	$N \geq 74$	5	$N \leq 11$	8	$N \leq 36$
LH_1	13	$N \leq 14$	10	$N \leq 11$	38	$N \geq 36$
LH_2	9	$N \leq 14$	10	$N \leq 11$	34	$N \geq 36$
LL_1	46	$N \geq 36$	6	$N \leq 11$	10	$N \leq 14$
LL_2	39	$N \geq 36$	8	$N \leq 11$	4	$N \leq 14$

Results

Analysis of Variance Related to the Posttest

The ANOVA using the total posttest measure as the dependent variable is displayed in Table 6. None of the F ratios for the two-factor treatment interactions or the one three-factor treatment interaction reached significance at the .05 level. A significant F ratio was obtained on each of the three main effects--frequency of concept moves, concept move interaction, and student abilities. The mean on the posttest for students experiencing the high frequency of concept moves was 35.5 (total possible was 68) compared to a mean of 32.4 for the students experiencing the low frequency of concept moves. A similar comparison for the two means of students experiencing the high and low levels of concept move interaction resulted in 35.9 and 31.9 respectively. Thus, the higher frequency of concept moves and the higher level of concept move interaction each facilitated student achievement on the total posttest administered the day following the completion of the instructional strategies.

Table 6

ANOVA Using the Total Posttest

Measure as the Dependent Variable

Source	df	Mean Square	F
(A) Frequencies of Concept Moves	1	186.834	4.223*
(B) Levels of Concept Move Interaction	1	311.911	7.051*
(C) Levels of Student Abilities	1	608.700	13.760*
A x B	1	.037	.007
A x C	1	117.085	2.647
B x C	1	50.354	1.138
A x B x C	1	1.959	.004
Error	69	44.238	

$^{*}\underline{p} \leq .05$, $\underline{F}(1,69) = 3.98$

The same analysis of variance procedures were used with each of the four cognitive subtests of the posttest as the dependent variable. The four ANOVAs for testing hypotheses 8-35 are summarized in Table 7. Significant main effects were recorded on the comprehension subtest but not on any other cognitive subtests. This finding may be attributed to the content information contained in the instructional strategies. Many of the concept moves were directed toward identifying examples and nonexamples of functions. Also, many of the items on the comprehension subtest required students to identify specific pairings as functions or nonfunctions. Hence, the instruction seemed more closely related to these items than any other subtest. Certainly this result supports the research hypotheses that more content information or more concept move interaction will increase student achievement.

A significant interaction was recorded in Table 7 on the application subtest between levels of concept move interaction and student ability levels. The B x C interaction is presented in Table 8 to demonstrate the interaction effect of student ability levels across the two levels of concept move interaction. An examination of the means in Table 8 indicated that the high level of concept move interaction facilitated student achievement for the high ability students but not for the low ability students.

Test items related to the composition of functions were included in the posttest even though the instructional strategies did not contain any information related to this topic. The students' performance on the composition of functions subtest can be interpreted as a measure of transfer. Table 9 contains the ANOVA using the composition of functions subtest as the dependent variable. A significant F ratio obtained in the analysis of the composition of functions subtest was attributed to the two levels of concept move interactions. In terms of an immediate transfer measure, the performance of students was facilitated by the higher level of concept move interactions on related concepts.

Table 7

Summary of ANOVAs Using the Cognitive Subtests as the Dependent Variable

Subtest		Knowledge		Comprehension		Application		Analysis	
Source	df	Mean Square	F	Mean Square	F	Mean Square	F	Mean Square	F
(A) Frequencies of Concept Moves	1	.007	.006	94.341	9.839*	24.904	2.291	5.461	1.031
(B) Levels of Concept Move Interaction	1	2.945	2.285	55.126	5.749*	27.393	2.520	4.[illegible]81	.808
(C) Levels of Students Abilities	1	1.337	1.037	91.525	9.546*	151.725	13.961*	.171	.032
A x B	1	1.121	.870	10.641	1.110	.811	.075	.237	.045
A x C	1	1.021	.792	21.845	2.278	15.170	1.396	4.181	.789
B x C	1	.033	.026	10.961	1.143	46.459	4.275*	2.660	.502
A x B x C	1	.705	.547	.171	.018	2.278	.021	1.602	.302
Error	69	1.289		9.958		10.868		5.300	

*$p \leq .05$, $F(1,69) = 3.98$

Table 8

B x C Interaction Table for the Application Subtest of the Posttest

	C_2	C_1
B_2	14.5	10.0
B_1	12.2	10.9

Note. B_1 and B_2 represent the low and high levels of concept move interaction, respectively. C_1 and C_2 represent ability levels I and II, respectively.

Table 9

ANOVA Using Composition of Functions Subtest as the Dependent Variable

Source	df	Mean Square	F
(A) Frequencies of Concept Moves	1	2.347	.99
(B) Levels of Concept Move Interaction	1	10.125	4.289*
(C) Levels of Student Abilities	1	6.125	2.594
A x B	1	.125	.053
A x C	1	7.014	2.972
B x C	1	6.125	2.595
A x B x C	1	.014	.006
Error	64	2.361	

*$p \leq .05$, $F(1,64) = 4.00$

Analysis of Variance Related to the Retention Test

The dependent variables for the ANOVAs testing hypotheses 36-70 were the total retention test measure and each of the four cognitive subtests of the retention test. The ANOVA using the total retention test measure as the dependent variable and testing hypotheses 36-42 is presented in Table 10. The only F ratio to reach significance was that of the classificational variable--student ability levels. Approaching the critical F value of 3.98 was the three-factor treatment interaction with an F ratio of 3.89.

Table 10

ANOVA Using the Total Retention Test

Measure as the Criterion

Source	df	Mean Square	F
(A) Frequencies of Concept Moves	1	131.546	2.492
(B) Levels of Concept Move Interaction	1	119.469	2.263
(C) Levels of Student Abilities	1	552.361	10.462*
A x B	1	42.177	.799
A x C	1	96.096	1.820
B x C	1	.778	.015
A x B x C	1	205.541	3.893
Error	70	52.795	

*$p \leq .05$, $F(1,70) = 3.98$

The four cognitive subtest measures of the retention test were each used as a dependent variable in analyses of variance. These four ANOVAs are displayed in Table 11. Four significant F ratios were recorded in the comprehension subtest ANOVA results--the three main effects and a two-factor interaction between frequencies of concept moves and student ability levels. The three-factor interaction effect using the application subtest and the knowledge subtest as the dependent variables reached significance. Another three-factor interaction effect approaching significance was recorded on the total retention measure. The only other significant F ratio in the four ANOVAs was produced by the student ability levels on the analysis subtest. Another main effect nearing significance on the application subtest was due to the levels of concept move interaction.

The three-factor interaction (A x B x C) results for the knowledge and application subtests are given in Table 12. A comparison of the effects due to the two experimental variables across the two ability levels is permitted using the data presented in this table.

A surprising result on both the knowledge and application subtests was found in a comparison of the student ability group means within the low frequency of concept moves strategy--the [illegible]B_2 and A_1B_1 cells of Table 12. It was surprising that there was no difference between the two ability groups (C_1 and C_2) when experiencing the low frequency of concept moves combined with a high level of concept move interaction. However, the largest difference between the two ability groups also occurred in the strategy containing a low frequency of concept moves--the LL instructional strategy, cell A_1B_1. Not surprising was the consistency across both tables of the highest and lowest means being recorded in the $A_2B_2C_2$ and the $A_1B_1C_1$ cells respectively of both tables. In other words the presence of both factors, the HH strategy, was more facilitative in promoting student achievement than the absence of both factors, the LL strategy. Another pattern is revealed by a comparison of the means across the two student ability levels in the A_2B_1 and A_1B_2 cells of both subtests. That is, the presence of just one main effect (A_2 or B_2) seems to be associated with little difference between the achievement of the two ability groups.

The two-factor interaction A x C effect for the comprehension subtest displayed in Table 13 is of interest due to the significant F ratios obtained on the main effects. A comparison of the means for the high ability level students reveals little difference between the means for the high and low frequencies of concept moves. However, a sizeable difference occurs for the lower ability students between the high and low frequencies of concept moves. Thus, the additional content information had more impact on the achievement of low ability students than on the achievement of the high ability students. One could hypothesize that additional content information beyond a certain point may not be increasingly facilitative for students. However, as the data in Table 13 indicate, this point of diminishing returns will be different for different ability levels.

Since the retention test was a parallel form of the posttest, it also contained the composition of functions subtest. Applying analysis of variance procedures to the composition of functions subtest revealed no significant F ratios. Thus, the facilitative transfer effect due to

Table 11

Summary of ANOVAs Using the Cognitive Level

Subtests as the Dependent Variables

Subtest		Knowledge		Comprehension		Application		Analysis	
Source	df	Mean Square	F	Mean Square	F	Mean Square	F	Mean Square	F
(A) Frequencies of Concept Moves	1	1.951	1.596	63.634	5.937*	15.472	1.397	3.378	.941
(B) Levels of Concept Move Interaction	1	.086	.071	45.917	4.284*	40.133	3.624	6.127	1.707
(c) Levels of Student Abilities	1	8.649	7.072*	101.701	9.489*	37.388	3.376	19.029	5.30[illegible]
A x B	1	.135	.111	.674	.063	26.361	2.380	3.289	.916
A x C	1	.005	.004	47.593	4.441*	.890	.080	4.138	1.152
B x C		.021	.018	.715	.063	.391	.035	.265	.074
A x B x C	1	5.195	4.248*	11.191	1.044	59.028	5.330*	1.059	.295
Error	70	1.223		10.717		11.074		3.590	

*$\underline{p} \leq .05$, $\underline{F}(1,70) = 3.98$

Table 12

A x B x C Interaction Tables for the Knowledge Subtest

and Application Subtest of the Retention Test

		Knowledge Subtest (6 items)		Application Subtest (22 items)	
		B_2	B_1	B_2	B_1
A_2	C_2	3.7	3.2	14.6	12.4
	C_1	3.2	3.0	11.5	13.1
A_1	C_2	2.9	3.3	13.3	12.3
	C_1	2.9	2.1	13.2	9.1

Note. A_1 and A_2 represent the low and high frequencies of concept moves, respectively.

B_1 and B_2 represent the low and high levels of concept move interaction, respectively.

C_1 and C_2 represent ability levels I and II, respectively.

the levels of concept move interactions found on the posttest was not present one month later.

Table 13

A x C Interaction Table for the Comprehension

Subtest of the Retention Test

	C_2	C_1
A_2	12.8	12.0
A_1	12.5	8.7

Note. A_1 and A_2 represent the low and high frequencies of concept moves, respectively. C_1 and C_2 represent ability levels I and II, respectively.

Discussion

Conclusions Related to the Frequencies of Concept Moves

The posttest data indicated that the students receiving a high frequency concept move strategy achieved significantly more than did the students receiving a low frequency concept move strategy. This finding supports our intuitive notion that there is a direct relationship between the amount of content information presented and the subsequent student learning. Somewhat disappointing was that the facilitative effect was not significant in the retention test results. An interesting question is suggested by the nature of mathematics and the results reported above --that is, in mathematics the content is sequential so if the two frequencies of concept moves were continued over a long period of time would there be a widening between the means of the two groups?

The posttest and retention test were partitioned into four cognitive level subtests--knowledge, comprehension, application, and analysis. A significant effect due to the frequency of concept moves was recorded on only one subtest of the posttest--the comprehension subtest. This finding indicates that additional content information in the form of concept moves improved students' understanding of the content contained in the instructional strategy. While a significant effect was also attributed to frequency of concept moves on the comprehension subtest of the retention test, there was a significant interaction between the frequency of concept moves and student ability levels that complicated conclusions based on this significant main effect. The interaction table for the frequency of concept moves x ability levels indicated that the high frequency of concept moves caused a substantial difference in student achievement for the ability level I students (lower ability students) but little difference in the achievement of the ability level II students. An examination of the group means of the low ability students indicated that the high frequency of concept moves facilitated the retention of content more than the low frequency of concept moves. The facilitative effect due to the high frequency of concept moves on the comprehension measure was not present in a comparison of the means of the high and low frequency concept move strategies using ability level II students. Thus, it appears that the high frequency of concept moves was more effective in the retention of content for the ability level I students than for the ability level II students.

Conclusions Related to the Levels of Concept Move Interaction

On the posttest the mean performance of the high level concept move interaction group was significantly greater than that of the low level concept move interaction group. This result confirms the belief held by many teachers and educators that student achievement is enhanced by student participation. However, the facilitative effect due to the high level concept move interaction strategy was not present in the total retention test results.

The comprehension subtest produced the only contrast of group means that was statistically significant on the cognitive level subtests of the posttest. One month later, on the retention test, the facilitative effects due to the high concept move interaction strategy was still significant on the comprehension subtest. Significant differences occurring on the comprehension subtest of the posttest and the retention test suggests that understanding and retention was greater for the students receiving the high level concept move interaction strategy than for the students receiving the low level concept move interaction strategy.

A significant interaction between the levels of concept move interaction and student ability levels occurred in the applications subtest of the posttest. The high level of concept move interaction facilitated achievement on the application subtest for the ability level II students, but for the ability level I students there was little difference in achievement on the application between the two concept move interaction groups. One explanation why the high level of concept move interaction improved achievement for ability level II students but not for the ability level I students depends on the student reactions to verbal participation. That is, the ability level I students may view the verbal interactions as "threatening" and become anxious about verbally participating, and this interferes with the learning process. The ability level II students were not anxious about the verbal interactions and were able to profit more from the teacher-pupil verbal interactions.

A significant three-factor interaction occurred on the knowledge and application subtests of the retention test. In both cases, the greatest mean performance was by the ability level II group receiving the high frequency of concept moves strategy and the high level of concept move verbal interaction. A consistent finding for both subtests was that the lowest mean performance was by those ability level I students who participated in the low frequency of concept moves and low concept move interaction strategy.

An interesting result is found in both interaction tables when the two ability level group means are compared across the four cells where either the high frequency of concept moves strategy or the high level of concept move interaction strategy (HL and LH), but not both, was employed. There seems to be no difference between the ability level I and II groups when experiencing either the high frequency of concept moves or the high level of concept move interaction. Disregarding the two ability levels, there seems to be little difference between the achievement of the HL and LH groups--that is, the high frequency of concept moves combined with the low level of concept move interactions (HL) and the low frequency of concept moves combined with the high level of concept move interaction (LH). One consistent and not unexpected finding in the knowledge and application interaction tables was that the HH strategy facilitated student achievement more than the LL strategy.

Implications for the Classroom

The final objective of research dealing with teaching strategies is to construct a theory which will guide the classroom teacher's behavior. The present study was an experiment to demonstrate that the amount of content information is under the teacher's control and is a significant factor affecting student achievement. However, a statistical significance does not automatically insure an educational significance. A comparison of the means of the high and low frequency concept move groups reveals a difference of 5% which is of questionable educational significance. However, a comparison of the same means for the comprehension subtest of the posttest and retention test indicates differences of 10% and 8% respectively. Certainly as the differences between teaching strategies approach 10% or higher, they become educationally significant. An examination of the interaction tables for the application and knowledge subtests of the retention test revealed differences in means between the HH and LL groups of 10% and 26% , respectively. Thus, in this limited example the HH strategy was educationally significant.

It is improper, from the results obtained in this investigation, to make an overall generalizaiton concerning the effects of a relatively high amount of teacher-pupil verbal interaction on student achievement. Certainly when significant differences due to the teacher-pupil interaction occurred, they favored the high level of concept move interactions. But an inspection of the significant treatment interaction tables revealed that this finding was not consistent across both ability levels and concept move levels. Thus, as a theory of instruction is established, it may prescribe different amounts of teacher-pupil verbal interaction for students of differing abilities.

A great deal of research has been directed towards comparing two extremes of pupil participation in the classroom dialogue. However, the results of this research are not conclusive--that is, it has not been clearly demonstrated that one extreme of teacher-pupil interaction is any better than the other extreme. Research has attempted to justify the use of more teacher-pupil verbal interaction on the basis of increasing student achievement, but the contradictory findings in the research literature do not permit this generalization. Perhaps the justification for more teacher-pupil verbal interaction will be based on something besides student performance on an achievement test.

Recommendations For Further Research

The literature is replete with statements confirming the nonexistence of a theory of instruction (Begle, 1973; Dodes, 1953; Gage, 1963). Although these statements, spanning the last two decades, indicate that the desired goal has not been attained, there has been progress. One area of progress has been concerned with the observation, recording, and/or describing of teachers' behavior in the classroom (Fey, 1969). Barr (1961) observed that there has been a shifting of emphasis from qualities of teachers to behaviors of teachers. The shift of emphasis was important because it is only through the use of reliable observational schemes that the behaviors of teachers can be manipulated and studied with controls similar to what might be expected in a laboratory setting. Hillway (1969) writes that the classroom may be made into a "de facto laboratory" if there are controls placed on the basic factors under consideration. Rosenshine and Furst (1973) echo the call for more research in classroom settings where the teaching act is monitored to insure fidelity to the treatment under investigation.

A desirable outgrowth of the proliferation of schemes for observing and describing the behavior of teachers is the increased variety of constructs to describe and monitor a teaching strategy. An analogy will demonstrate how a variety of constructs can be useful in research concerned with teaching strategies. In order to evaluate the effect of some treatment on students' behavior, a variety of measures can be taken to pinpoint the treatment effects. Some measures used in the past are different cognitive levels, attitudinal changes, transfer measures, and retention measures. In fact, Wilson (1971) stated that using just one measure may lead to incorrect conclusions. Applying a similar logic to the descriptions and monitoring of teaching strategies would suggest the use of a variety of different constructs. Two advantages that may be realized from using a variety of constructs are (a) more precise defining of teaching strategies and (b) more detailed monitoring of the teaching strategy. Realization of these advantages would in turn enable researchers to replicate teaching strategies and also compare results between separate studies. Some of the constructs for describing and/or monitoring of teaching strategies will be discussed in the following paragraphs.

A taxonomy of moves for the teaching of concepts was developed by Henderson (1967). These concept moves have been used as the basis for defining teaching strategies as a sequence of concept moves--usually presented to students in a programmed format. Most of these programmed teaching strategies have investigated the relative efficacy of different sequences of concept moves. However, another use of concept moves is to quantify the amount of content information (e.g., the number of concept moves) in a teaching strategy. Certainly in a programmed format the number and sequence of concept moves are easy to obtain; however, through the use of tape recorders, this same information can be obtained from the classroom discourse. In addition to comparing teaching strategies in terms of the number or the sequence of concept moves, another comparison of teaching strategies is permitted by considering the variety of concept moves they contain. Perhaps the "richness," i.e., variety of moves of a teaching strategy, is a significant factor affecting students' understanding. Suppose that of two teachers, A and B, teacher A consistently uses a greater variety of concept moves than teacher B. Would this finding imply (a) that teacher A was more creative or (b) that teacher A possesses more knowledge on the content under consideration, or (c) something else altogether?

Using the notion of concept moves as a unit of information, it is possible to describe the amount and sequence of information in a teaching strategy. Another use of the concept move might involve a time element, in other words, the "pacing" of content could be described by the frequency of concept moves over a period of time. One could then ask two questions: How should the pacing change for different ability levels? How should the pacing change for different types of concepts?

Teaching strategies defined in terms of concept moves and implemented in a near natural classroom setting are important for two reasons. First, mathematics teachers spend a great deal of time engaged in the teaching of concepts. Second, the research literature reveals that there is a paucity of studies investigating how the cognitive aspects of the classroom dialogue are related to student learning. Some examples of the constructs available for investigating the cognitive aspects of teaching strategies are:

(a) ventures, episodes, and monologues (Smith et al., 1967);

(b) deductive, inductive, classifying, and analyzing categories (Cooney & Henderson, 1972); and

(c) substantive-logical meaning processes (Bellack, Kliebard, Hyman, & Smith 1966).

Another consideration in the description of a teaching strategy is related to the type of content contained in the strategy. Cooney, Davis, and Henderson (1975) describe three types of knowledge, concepts, generalizations, and skills, each of which may be used as the focus of a teaching strategy. Cooney et al. have developed a taxonomy of moves for the teaching of generalizations. They also discuss how the generalization moves can be used to differentiate between a guided discovery strategy and an expository strategy. In addition to studying teaching strategies where the number or sequence of moves for the teaching of generalizations is manipulated, it may prove fruitful to investigate combinations of generalization moves and concept moves. Concepts form a necessary part of generalizations, yet the study of teaching strategies containing both types of moves represents an unexplored approach.

It would be interesting to compare the relative importance that teachers place on teaching each of these types of knowledge. All three types--concepts, generalizations, and skills--are important. Teaching strategies with different objectives (e.g., attainment of concepts, application of generalizations, or improvement of skills) must be defined and investigated to determine how a teacher should behave relative to each type of objective.

References

Barr, A. S., Worchester, D. A., Abell, A., Beecher, C., Jensen, L. E., Pesonto, A. L., Rigness, T. A., & Schmid, J., Jr. Teacher effectiveness and its correlates. Wisconsin studies of measurement and prediction of teacher effectiveness: A summary of investigations. Madison: Dembar Publications, 1961.

Begle, E. G. Some lessons learned by SMSG. The Mathematics Teacher, 1973, 66, 207-214.

Bellack, A. A. The language of the classroom. U. S. O. E. Cooperative Research Project No. 2023. New York: Columbia University Teachers College Institute of Psychological Research, 1965.

Bellack, A. A., Kliebard, H. M., Hyman, R. T., & Smith, F. L., Jr. The language of the classroom. New York: Teachers College Press, 1966.

Cooney, T. J., Davis, E. J., & Henderson, K. B. Dynamics of teaching secondary school mathematics. Boston: Houghton-Mifflin, 1975.

Cooney, T. J., & Henderson, K. B. Ways mathematics teachers help students organize knowledge. Journal for Research in Mathematics Education, 1972, 3, 21-31.

Dodes, I. A. The science of teaching mathematics. The Mathematics Teacher, 1953, 46, 157-166.

Fey, J. T. Classroom teaching of mathematics. In J. W. Wilson & L. R. Carry (Eds.), Reviews of recent research in mathematics education. Studies in mathematics (Vol. 19). Stanford, Calif.: School Mathematics Study Group, 1969.

Gage, N. L. Paradigms for research on teaching. In N. L. Gage (Ed.), Handbook of research on teaching. Chicago: Rand McNally, 1963.

Gage, N. L. Teacher effectiveness and teacher education. Palo Alto, Calif.: Pacific Books, 1972.

Henderson, K. B. A model for teaching mathematical concepts. The Mathematics Teacher, 1967, 60, 573-577.

Hillway, T. Handbook of educational research. Boston: Houghton-Mifflin, 1969.

Hughes, M. M. Utah study of the assessment of teaching. In A. A. Bellack (Ed.), Theory and research in teaching. New York: Columbia Teachers College Press, 1963.

Meux, M., & Smith, B. O. Logical dimensions of teaching behavior. In B. J. Biddle & W. J. Elena (Eds.), Contemporary research on teacher effectiveness. New York: Holt, Rinehart & Winston, 1964.

Rosenshine, B., & Furst, N. F. The use of direct observation to study teaching. In R. M. W. Travers (Ed.), Second handbook of research on teaching. Chicago: Rand McNally, 1973.

Smith, B. O., Meux, M., Coombs, J., Nuthall, G., & Precians, R. A study of the strategies of teaching. Urbana: Bureau of Educational Research, College of Education, University of Illinois, 1967.

Snow, R. W. (Ed.). A symposium on heuristic teaching (Technical Report No. 18). Stanford, California: Stanford Center for Research and Development in Teaching, 1970.

Swank, E. W. A comparison of selected strategies for the teaching of mathematical concepts (Doctoral dissertation, University of Georgia, 1973). Dissertation Abstracts International, 1974, 34A, 5001. (University Microfilms No. 74-4887)

Thomas, H. L. The concept of function. In M. F. Rosskopf (Ed.), Children's Mathematical Concepts: Six Piagetian studies in mathematics education. New York: Teachers College Press, 1975.

Wilson, J. W. Evaluation of learning in secondary school mathematics. In B. S. Bloom, J. T. Hastings, & G. P. Madaus (Eds.), Handbook on formative and summative evaluation of student learning. New York: McGraw-Hill, 1971.

Winer, B. J. Statistical principles in experimental design. New York: McGraw-Hill, 1962.

Design Problems in Research on Teaching Strategies in Mathematics

Richard L. Turner

Indiana University

This paper focuses on two sets of design problems in research on teaching strategies in mathematics, grades 1-12. The first set involves the definition and organization of domains of variables believed to be relevant to mathematics instruction, so that samples of variables drawn from these domains yield as much information as possible about them. The second set of problems involves the practical matter of relating multiple classes of variables to each other in such a way that reasonable inferences about their relationships can be drawn.

Relative to these points, a critical consideration is that the various investigations conducted in this research domain yield evidence which may be systematically incorporated into arguments about the true relationships in it. For example, the central claim inherent in the taxonomical work on teaching by Smith, Meux, Coombs, Nuthall, and Precians (1967), Henderson (1972), and Cooney (1974), is that the "moves" made by teachers, and the order of these moves ("strategies"), will ultimately be found to have a significant bearing on the acquisition of mathematics concepts, principles, and skills among students. No single piece of research, experimental or otherwise, will prove or disprove the arguments surrounding this claim. A set of properly conceived investigations should, however, yield evidence that the claim is or is not empirically valid. If it is not valid, this evidence should suggest alternative arguments which seem best to account for the relationships found in the domain. The point to be made is that the results of individual research projects focused on teaching strategies are not ends in themselves. Rather the projects are used to support arguments about the true composition of the domain, and through these arguments they attain a generality far beyond that attainable within any individual investigation.

Defining and Sampling the Relevant Domains

The first step in a properly conceived research design is to identify the domains of variables to be covered in the design. There are four such domains in research on teaching strategies in mathematics. The first is the substantive domain of mathematics as taught in elementary and secondary school. According to Henderson, this domain may be divided into sets to include concepts, principles, and skills. The sets may be further divided into subsets or classes. The second domain is comprised of teacher actions and interactions believed to be relevant to the acquisition by students of the substance of the first domain. The relevant domain of teacher action has two sets in it: "moves" and "strategies" (combinations of moves). Each of these sets, according to both Smith et al. (1967) and Henderson (1972), may be further subdivided. The third relevant domain is comprised of student attributes which, in addition to the actions of the teacher, are believed to be correlated with the acquisition of mathematics concepts, principles, and skills. There are numerous classes of student attributes, but not all are equally salient in the acquisition of mathematics content. The fourth relevant domain is comprised of those performances by students which are taken as indicators that particular concepts, principles, or skills in mathematics have been acquired. As in the other domains, the set of all indicators of mathematics attainment can be divided into subsets, with each resulting set theoretically pointing to different levels of and/or different types of mathematics attainment.

In addition to the four domains of variables noted above, a fifth domain which may be labeled "setting variables" may be also recognized as generally pertinent to research on teaching. This domain is interpreted here to include the type of school organization and "climate" in which a person teaches, the attitudes of the community toward schooling, the types of materials provided for instruction, and the like. For present purposes the variables in this domain will be treated as "givens" and will be taken into account only in a very general way (for example, in extracting between-school or between-district variance in the event a particular investigation is subject to error through the intrusion of variables from this "setting" domain).

As noted above, each of the domains is composed of subsets or classes of variables. For example, mathematics concepts and skills are viewed as different classes of variables, and so are moves and strategies, even though strategies are comprised of combinations of moves. More generally, if one conceptualized the domains as dimensions, each dimension may be further conceptualized as having an array of classes along it. There are thus four arrays of classes of variables. Within each dimension, the classes of variables hold either analytic (or definitional) relationships, or else empirical correlative relationships, to each other. The relationships between dimensions, on the other hand, are all ordered conditional ones. These relationships may be expressed in general hypothetical form as follows: If student attribute $\underline{A}$ is divided into levels $\underline{L_1}, \ldots, \underline{L_n}$, and if concept $\underline{C}$ is to be taught (is the teachable object), then teaching strategies $\underline{S_1}$; $\underline{S_2}$, $\underline{S_3}$ will produce student outcomes $\underline{0}$, such that for all indicators of $\underline{0}$, $\underline{I_1}$, $\underline{I_2}, \ldots, \underline{I_n}$, $\underline{S_1 0} > \underline{S_2 0} > \underline{S_3 0}$, for all levels $\underline{L_1}$, $\underline{L_2}, \ldots, \underline{L_n}$ of $\underline{A}$.

Buried within this generalized proposition are two important tasks to be performed by researchers who wish to make empirical tests of the proposition cast in specific form, i.e., when actual variables are substituted for the general terms. The first of these tasks is to attempt to represent in each specific term chosen as many elements of the class to which it belongs as possible. Thus, each specific variable or element selected may be viewed as a sample of a class or set to which generalization might be made. The second task is to increase the potential precision of the experiment, quasi-experiment, or survey to the greatest degree. The reason for attending to precision is that the power of a specific teaching move or strategy is never certain. One thus wishes to optimize, insofar as he can, the chances that the true results will appear by systematically taking out known sources of variation in the outcomes sought. The importance of these two tasks varies with each of the four domains under consideration, thus each task needs to be examined relative to each domain.

Domain 1. Mathematics Concepts, Principles, and Skills

A major portion of the work done by Henderson (1970, 1972) and presented by Cooney (1974) has been to develop a taxonomy of teachable objects in mathematics. The structure and components of this taxonomy are summarized in Figure 1 and some symbols added to them.

As may be noted in Figure 1, each class of the taxonomy has many elements in it. For example, the class of denotative concrete singular concepts (Dcs) has in it all such concepts in mathematics grades 1-12, or Dcs_n concepts, where n must be a very large number. The same must be true of the other classes in the taxonomy, excluding nondenotative concepts ($\sim$ D), which is regarded as mathematically uninteresting.

From the viewpoint of the researcher, the existence of large classes of teachable objects of unknown heterogeneity poses a substantial problem. To avoid conducting an indefinite number of experiments involving every conceivable teachable object, he must select for each experiment those teachable objects from which generalization can be made to the members of the homogeneous class to which the object chosen belongs. The problem is to increase the probability that the classes of teachable objects are homogeneous before conducting the experiment. The Henderson taxonomy does not address this problem, but it should nonetheless be dealt with to enable systematic research to occur in this field. Following are some steps that might be taken to resolve this problem.

First, beginning with the classes already existing in the Henderson taxonomy, each class should be stratified according to the general grade level (or developmental level in mathematics) at which each particular concept, principle, or skill is typically taught. Thus, the singular concrete denotative concepts in class Dcs would be ordered to perhaps six strata or levels. Second, within each stratum, the concepts, principles,

	Concepts					Principles		Skills
	Denotative		Attributive		∼D	Generalizations	Prescriptions	
	Concrete	Abstract	Concrete	Abstract				
Singular	Dcs_1 . . . Dcs_n	Das_1 . . . Das_n	Acs_1 . . . Acs_n	Aas_1 . . . Aas_n		G_1 . . . G_n	P_1 . . . P_n	S_1 . . . S_n
General	Dcg_1 . . . Dcg_n	Dag_1 . . . Dag_n	Acg_1 . . . Acg_n	Aag_1 . . . Aag_n				

Figure 1. A representation of the taxonomy of teachable objects in mathematics.

or skills should be redivided according to how easy or difficult they are for students to learn (i.e., for teachers to teach successfully). The reasons for taking these steps are: First, mathematics content is already graded so that an ordered series of concepts, principles, and skills appear in it, and any experiment is almost certain to be done within a particular level of this series. Second, the ease-difficulty of a teachable object is correlated (but imperfectly so) to its placement in the series. Third, separating teachable objects according to ease-difficulty makes possible the selection of the more difficult objects for the experiment to be conducted. Selecting difficult concepts, principles, or skills is desirable because "easy" ones are by definition those that can be learned about equally well under any teaching strategy. Thus, the power or superiority of one strategy versus another would not be expected to appear unless "difficult learnings" are involved in the experiment. Although these first two steps seem formidable, the information on which they are based should already be available in the literature of mathematics education and should be easy to retrieve.

A third step in creating more homogeneous classes of teachable objects seems more difficult than the other two. It requires the identification of substantively related families of concepts, principles, or skills across the series of strata developed in mathematics content in step 1. These families, for example, might be comprised of concepts with common elements such as common fractions, proportions, decimals, and percents. If such families do exist extensively in mathematics, identification of them and stratification according to them is important insofar as one would expect teaching strategies successful in attaining superior outcomes in one branch of the family to have better than random probability of generalizing to other branches of the family.

If the three steps described above were to be taken in the domain of teachable objects, the result would be a set of sampling frames within each subdomain of concepts, principles, and skills. For example, in the domain of concepts, and in the set of denotative concepts, a minimum of 12 sampling frames might be foreseen: six strata or developmental levels of concepts and within each, two levels easy and difficult. Finally, an additional stratification by families of substantively related concepts might be possible. It seems likely, however, that stratification by these families is not perfectly replicated or balanced across all the twelve frames, so that some imbalance would result, i.e., not all families can be represented in all frames.

Once these frames are developed, the stage is set for sampling from them. The exact method of drawing the sample is, however, a practical matter of research design and is discussed in Section II.

Domain 2. Student Attributes

The reasons for attending to student attributes in research designs involving teaching strategies are quite different than those for developing sampling frames and more homogeneous classes in the universe of teachable objects in mathematics. Student attributes are employed, on the one hand, to increase the statistical precision of the design, and on the other, to provide the possibility of obtaining greater information about the effects of particular teaching strategies on different types of students. Increases in the precision of designs come about by extracting from the variance of the indicators (I_1, I_2,...,I_n) of the student learning outcomes those portions which are attributable to the antecedently measured attributes of the students. This helps to minimize the residual or error variance. By minimizing the error variance, the probability of showing effects for teaching strategies may be increased.

A general design strategy which seeks to increase precision also may be employed frequently to increase the amount of information extracted from the research, depending on the specific design used. In designs employing analysis of covariance, randomized blocks (matched-randomized), or residual gains analysis, the variance attributable to the antecendently measured student attribute(s) is usually discarded as uninteresting, i.e., no specific information is extracted from it. In stratified designs, on the other hand, about the same level of precision can be attained, but the otherwise discarded variance can be assigned to levels of a student antecedent attribute (L_1, L_2,...,L_n of A), so that effects of the various teaching strategies on different types of levels of students may be extracted from the design. Thus, the information available from the design is potentially increased.

The practical problem confronted by the researcher is not simply which type of design to use, but more importantly, which student attribute(s) to assess as the antecedent. In research on teaching strategies, one important criterion to use in the selection of this attribute is that measures of it be readily available. If they are not, and different levels of the attribute are subsequently shown to relate to different types of teaching strategies, teachers will be unable to act on this information since they will be unable to stratify their students and differentiate strategies for students in these strata. A second important criterion for attribute selection is the expected correlation between the antecedent measures and the learning outcome measures. If this correlation is low, little gain in precision will occur, and, unless there are interaction effects between the levels of the antecedent measures and the outcome measures, little will be gained from introducing the attribute into the design.

The student attributes which best fit these criteria are, theoretically, specific aptitudes for learning mathematics in its various branches. Quantitative estimates or measures of these aptitudes can be generated either on the basis of the past performance of the student in that branch (e.g., arithmetic, algebra, geometry) or by a test which covers concepts, principles, or skills which the student would be expected to have learned. The notion that other attributes of students such as attitudes toward mathematics or personality characteristics should be employed in research designs involving teaching strategies is probably unsound. The influence of such variables on mathematics learning is not well established. Moreover, any significant influence from these sources would be incorporated into the variance of aptitudinal measures since the latter would reflect any set of attributes consistently correlated with previous attainment in mathematics.

Domain 3. Teaching Moves and Strategies

This domain confronts the researcher with a critical choice at the very outset. This choice is between doing experiments which test, as the experimental treatments, selected teaching strategies versus teaching preservice or inservice teachers the relevant array of moves for each type of teachable object, then letting them develop the teaching strategies as they see fit. Up to this writing, the experiments conducted have chosen the first of the two alternatives. A good case can be made, however that this alternative has distinct difficulties associated with it.

First, a strategy is a combination of moves. To say that it is a "temporal sequence of moves" does not add information since two or more moves cannot occur simultaneously. Viewed as a combination of moves, it is intuitively clear that there are in excess of n factorial strategies available for each set of n relevant moves, since a move may be repeated in a strategy. For example, there are six relevant moves available for teaching denotative concepts. Since two or more moves must occur in sequence for a strategy to occur, and since each strategy has a fixed order of moves, there are at least 6! strategies for teaching denotative concepts. Since moves can be repeated, there are of course a great many more than 6! strategies possible. Given the fact that different sets of moves are available for teaching concepts, principles, and skills, it is quite clear that the number of strategies that might be experimented with is very great.

The difficulty confronted by the researcher is, of course, to select from among the possible strategies that one or that set with which he wishes to experiment. Hypothetically, the researcher needs both to select the most powerful strategy for teaching a given concept and to be able to generalize from the strategy he has chosen to some subset of similar strategies. The procedures he is to use to make these choices, other than trial and error, are not at present altogether clear. In existing research, the choices appear to be made intuitively and left unexplained.

A second difficulty with a research approach which tests the efficacy of particular strategies lies in the utility of the outcomes for teachers. To see this difficulty one must first hold in mind that there is an n-dimensional matrix of teaching strategies for concepts, principles, and skills, and there is also an n-dimensional matrix of teachable objects.

If different strategies are needed to teach the different classes of teachable objects (the subclasses of concepts, principles, and skills), the total number of strategies a teacher would need to learn and be able to use would create a skill acquisition and memory load of immense magnitude -- one probably impossible to deal with.

Although these difficulties are perhaps not insurmountable, they do suggest that an alternative kind of research on teaching moves and strategies might be undertaken. In simplest terms, the alternative research strategy is to train teachers to become skillful in making specific teaching moves and to familiarize them with the concept of teaching strategies, but not to attempt to train them on specific strategies. Rather, each teacher would be left to generate strategies situationally. Aside from the fact that this approach reduces the memory load for the teacher, an important reason for considering it lies in the distinct possibility that the course of moves in a strategy actually depends on the feedback the teacher is receiving from students. Indeed, the original work by Smith et al. (1967) and Smith and Meux (1970) derives teaching moves as monadic units of teacher behavior out of dyadic (or interdependent, response-contingent) teacher-pupil classroom behavior. In this work, a strategy was a series of maneuvers on the part of a teacher toward a particular teaching goal. These maneuvers were not completely independent of what the participating students did. Rather, the course of the moves depended to some degree, but not entirely, on how one or more students responded to each successive move.

If the notion that student response following a move is incorporated as a necessary attribute into the concept of teaching strategies, the definition of the latter changes and becomes dyadic as follows:

> A teaching strategy is a combination of moves the exact course of which depends on the response of the student(s) following each move which solicits a response from students. (dyadic definition)

This is in contrast to:

> A teaching strategy is a combination of moves. (monadic definition)

The difference between these two definitions from the researcher's viewpoint is that the monadic definition points the stream of research toward programmed instruction or "programmed teachers" in which the order of moves in a strategy is fixed. It is notable that Nuthall's dissertation (1967), the first empirical research in this area, used programmed materials as the treatments and apparently greatly influenced the course of subsequent research. The alternative stream of research, that based on the dyadic definition of a strategy was, however, never developed, or if developed, never published. The writer suggests that it be developed as the alternative main stream of research on teaching moves and strategies. The two main streams of research, one based on the monadic definition and the other on the dyadic definition are outlined below.

Research based on a monadic concept of teaching strategies. As indicated in earlier paragraphs, a central problem in research in this stream is how to sample strategies. This sampling should be done in such a way that those strategies selected are: (a) potentially the most powerful ones, (b) represent a class of similar strategies, and (c) are sufficiently few, or are sufficiently easily learned, to be incorporated into the skills of teachers without creating a cognitive overload.

One way to approach this sampling problem is to begin not with the strategies themselves, but with an assumption about the manner in which teachers hold or remember pedagogical knowledge. This assumption is that teachers adopt approaches to or models of teaching under which they then subsume clusters of knowledges and skills. This assumption seems implicit in Henderson's taxonomy (1970) and is used by Cooney (1974, p. 161) and by Lester (1974) in discussing the skills of mathematics teachers. Under it, approaches varying from highly didactic or "deductive" such as exposition or rule-example through advanced organizer (analogy), guided discovery, and discovery or "inductive" approaches are selected as the first step. Second, strategies are grouped under these approaches according to the degree to which they exemplify the approach and the extent to which they are suggested by or supported by the research literature. For example, the "ruleg" or rule-example approach is exemplified by strategies which open with one or more connotative moves followed by denotative moves. With respect to the denotative moves, some of the recent literature on concept learning (Markle & Tieman, 1971; Thiagarajan, 1971) suggests that whether or not the extended examples and nonexamples given are matched or not matched to those initially given is an important aspect of learning and retaining concepts under a "rule-example" approach.

The third step is to group the strategies according to the number of moves which must be made clear in order to give the student extensive information about the concept, principle, or skill. Hypothetically, the most powerful strategies will be those that provide the most information in the fewest moves. This hypothesis might be false, however, so that sampling strategies from different groups ordered according to the information given per move is an important consideration in actually conducting an experiment.

The intent of these three steps is to organize each domain of strategies into sampling frames. From these sampling frames, the most promising strategies can then be sampled for experimental work and, if found promising, for teaching to teachers.

Research based on a dyadic concept of teaching strategies. The general objective of this stream of research is to determine whether or not teachers trained to make selected teaching moves smoothly and accurately obtain, under experimental conditions, better learning outcomes from students than teachers not thus trained. A key factor in this stream is the selection of those moves on which teachers are to be trained, since within existing time constraints all experimentally trained teachers probably cannot be trained to make all the moves in all of the domains. Although individual researchers may have their own strategies for approaching the selection problem, a good way to approach it appears to be to follow the taxonomy of moves for teachable objects already produced by Cooney, Davis, and Henderson (1975). Thus, an initial step would be to restrict the moves to be learned to a specific subset such as those related to denotative concepts. The second step would be to make a division between the connotative and denotative moves. A third step, in the judgment of the writer, would be to regroup the moves within these subsets according to the subjectively estimated probability that a particular move will have to be made by a teacher instructing at a particular level if the concept (principle or skill) is to be learned by students. Thus, teachers in the primary grades will be taught a somewhat different set of moves than those teaching algebra or geometry in secondary school.

Following these steps focused on the selection of the concepts on which the experimental teachers are to be trained, the next step is to produce the procedures by which the teachers are actually to be trained. These procedures are dealt with in detail in the paper by David Gliessman, and will not be elaborated in this paper. The final step prior to actually doing an experiment is then to train the teachers to criterion, with the specific moves on which they are trained determined by the arrangement of the experimental treatments. For example, one group might be trained entirely on giving examples and nonexamples, another on sufficient conditions and differentiation moves and on example-nonexample moves, and so on, until the experimental treatments the investigator wishes to employ are exhausted. The number of treatments the experimenter seeks to employ must of course be limited since each addition increases the complexity of the experimental design, which, given the other variables to be taken into account, is already complex.

A feature of this stream of research which must occur during the experiment is the observation of samples of the actual moves made by the experimental teachers during instruction. These observations must be done in order to determine whether or not the differences in training the teachers appeared in their actual behavior. Additionally, these observations may be correlated with student learning outcomes to discover associations between the latter and the teacher behaviors actually used by the trained teachers. A simple model showing the relationships in question appears in Figure 2.

In this model, relationship 1 is between the observed behaviors of the trained groups (X_A vs. X_B vs. X_C); relationship 2 is comprised of the correlations of the pooled behaviors X_A, X_B, X_C and the pooled outcomes O_A, O_B, O_C, while relationship 3 is the contrast in outcomes O_A vs. O_B vs. O_C, for trained teacher groups A, B and C, when relationships 1 and 2 are disregarded.

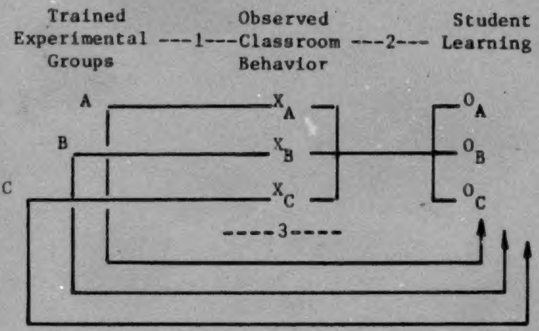


Figure 2. Relationships among training, classroom behavior, and student learning.

Domain 4. Student Learning Outcomes

As a rule, one gets what he teaches for. Thus, if students are taught applications, they will perform when applications are requested, or if they are taught to verbalize generalizations, they will return these to the teacher when tested for them. It is always difficult for researchers to keep this principle in mind since they are invariably hopeful that students will generalize their learning considerably beyond what was actually taught.

The critical tasks for researchers in considering learning outcomes associated with teachable objects in mathematics are first, to set forth the indicators in student performance which one will take as evidence of learning, and second, to be able to specify, if at all possible, the relationship between the experimental treatments and different types or levels of indicators. In an ideal taxonomy of teachable objects x teaching strategies, the performances which indicate that students grasp a concept or principle, or can perform a skill would be given. In the Henderson taxonomy as presented by Cooney (1974), however, these indicators are not developed, or at least not clearly developed. This gap in the taxonomy throws to the research investigator the problem of completing the taxonomy on an empirical basis.

To see this point more clearly, it is useful to return to the general proposition stated earlier. In this proposition the claim is made that "teaching strategies $\underline{S}_1$, $\underline{S}_2$, $\underline{S}_3$ will produce student outcomes $\underline{0}$, such that for all indicators of $\underline{0}$, I_1, $I_2, \ldots, \underline{I}_n$, $\underline{S}_1\underline{0} > \underline{S}_2\underline{0} > \underline{S}_3\underline{0} \ldots$" It is clear, however, that this proposition can be true only if $\underline{I}_1$, $\underline{I}_2, \ldots \underline{I}_n$ are a homogeneous set. If the set of I is heterogeneous, then it could be true that $\underline{S}_1\underline{I}_1 = \underline{S}_2\underline{I}_1 = \underline{S}_3\underline{I}_1$ but that $\underline{S}_1\underline{I}_2 > \underline{S}_2\underline{I}_2 > \underline{S}_3\underline{I}_2$, and that $\underline{S}_1\underline{I}_3 > \underline{S}_2\underline{I}_3 > \underline{S}_3\underline{I}_3$. Indeed, this appears to be what happened in the Dossey-Henderson study as reported by Cooney (1974, p. 169). When it does happen, the investigator reports the differential outcome and in essence completes the taxonomy for the set of strategies tested x the teachable object chosen x the outcomes obtained, according to the indicators used, for the level of student employed in the experiment. The essential point to be made here is that the taxonomy cannot be accurately completed, which is the objective of the research, unless a spectrum of indicators of concept, principle, or skill learning is initially provided by the investigator.

What the necessary and sufficient (or at least sufficient) indicators are for concept, principle, and skill learning in mathematics is a matter better dealt with by mathematicians than by the present writer. A few observations about indicators might, nonetheless, be made at this point.

First, one way to develop indicators is to use the teacher moves in the taxonomy as the foundation for the sets of indicators. This can be done because each teacher move displays a different aspect of knowledge about the teachable object. Thus, differentiation moves or instancing moves are implicitly taken as evidence of knowledge of the object. Collaterally, student responses to questions based on these moves are evidence that he or she has attained the knowledge. Technically speaking, moves should be organized in such a way that as one goes from one move to the next successively more information about the state of knowledge of the learner is revealed. For example, a differentiation move seems to be at a slightly higher level of knowledge or learning than a positive example move since differentiation requires knowledge of the criterial attributes of two concepts rather than only one concept.

If teaching moves are used as the foundation for generating indicators of student learning, one should notice that the second of the two problems initially stated is helped toward solution. Namely, an implicit relationship is formed between what the teacher does (make moves) and what the student is tested over (items which require the knowledge revealed by these moves). Indeed, it is perfectly reasonable to hypothesize that the sequence of moves made by the teacher is less important to learning outcomes than whether the combination of moves made increases the total information (or knowledge) available to the student.

Second, although it is true that problem-solving or "application" is an extremely important outcome of school mathematics, it is also true that what is learned at one level of mathematics must facilitate what is learned at subsequent levels if mathematics instruction is to succeed in the long run. It follows that while one very much needs to develop indicators of the student's ability to solve problems involving a concept, principle, or skill, it is also important to develop indicators that the learning associated with a teaching strategy results in proactive positive transfer to other related mathematical learnings. To develop indicators of this type requires the experimenter to provide two successive learning tasks for students. In the first, students are taught a teachable object by the several strategies selected by the experimenter. The outcomes can be tested by the usual indicators (e.g., items representing levels of Bloom's taxonomy). In the second experiment, however, the learning task involving a new teachable object is presented to all experimental groups by the same strategy, and either the time needed to acquire mastery of the new teachable object, or the degree of mastery of it, or both, are taken as the criterion variables. If one of the strategies utilized in the first experiment has a facilitative effect on the acquisition of the new teachable object, a between-groups comparison of the level of performance on the criterion variables in the second set will reveal the degree of facilitation the earlier teaching and learning had on the later learning.

Designs

The objective of individual experiments examining strategies in teaching mathematics is to obtain an estimate of the truth of the underlying arguments in the simplest possible ways. Central among these arguments is that the combination of types of moves made by the teacher is significant to learning teachable objects in mathematics. A major feature of teaching moves, in addition to the combination of types of moves, is the quality of each move made. The quality of a move has at least two attributes: clarity and the information revealed by the content of the move. Dodd (1974) has recently shown that the rated clarity (clarity as a high inference variable) of the teacher's presentation in teaching mathematics (fractions) accounts for a very substantial percentage of the variance of learning attributable to teaching. Thiagarajan (1971) and Markle and Tieman (1971) have shown, in concept learning, that the convergence-divergence (or degree of matching) of examples and nonexamples is an important aspect of the content of exemplification moves. It follows that in the design of experiments involving strategies for teaching concepts, a warranted inference relative to the comparative effects of two or more strategies might be drawn only if each strategy involved is equivalent in clarity and in the content of the exemplification moves. Otherwise, the alternative hypothesis that any differences among strategies is attributable to differences in either clarity or in content is viable.

Controlling for or directly testing for the influences of the quality of teaching moves is at present critical for the survival of research on teaching strategies. The review of research on strategies by Dossey (in this monograph) produces relatively little evidence to increase confidence that strategies make any difference. In the studies reviewed, however, little control over the quality of the moves was apparently exercised. Thus, an initial move to be made is to produce a design which is similar to those previously used, but either (a) closely controls quality or (b) examines variations in quality while testing differences in strategies at the same time. Of these two alternatives, (b) is initially the better, although not necessarily the simpler, since it can be made to yield more information than alternative (a).

Design 1.1 Monadic Definition of "Strategy"

Intent. The intent of this design is to determine the contribution of (a) two teaching strategies, (b) two levels of clarity, and (c) three levels of convergence-divergence of examples to selected pupil learning outcomes relative to one, and possibly two or more, denotative concepts. This basic intent, however, may be intrinsically compromised in the design since it is constructed as if clarity, exemplification, convergence-divergence, and teaching strategies are independent or orthogonal variables. Quite clearly, they could be correlated. If they are, the outcome could be (a) immense difficulty developing the treatment (for example, making up a treatment that is low in clarity but has several divergent examples) or (b) weak or absent main effects and large difficult to interpret interactions.

Design validity. The emphasis in the design is on internal rather than external validity, thus sampling from the various relevant universes is nonrandom or "fixed." It follows that generalization to different levels of students, to the type of teaching moves, and to the universe of teachable objects cannot be made. An important consideration in the design, however, is that it is to be replicable at different ages or grade levels and with different teachable objects in the concept domain. On the whole, independent replications of the design in this way will build (or fail to build) confidence in the generality of the effects more quickly than will any alternative procedure.

An aspect of the internal validity of the design which cannot be overlooked is the ease-difficulty of the teachable objects chosen. To make a fair trail of the strategies argument or of clarity or the content of the examples, the teachable objects must be drawn from those identified as "difficult." A problem for the experimenter is to find a set of procedures by which to isolate the difficult concepts. One procedure, mentioned in Section I, is to refer to the research literature. A second, and probably more expedient procedure, is to have mathematics educators or mathematics teachers familiar with the concepts taught at a particular level (or narrow range of levels) rate concepts for how easy or hard they are to teach and correlatively, how easy or hard they are for students to learn. The concepts used in the design would then be selected on the basis of these ratings, with the additional constraint for younger students that they be within the range of normal development for the ages of the children.

In addition to the ease-difficulty matter, the normal items of internal validity are to be observed. These items include the random assignment (but not necessarily random selection) of subjects, the effects of maturation, attrition from the experiment, the intervention of extraneous factors during the experiment, and so on.

Treatment construction. An initial choice is whether to use one concept or more than one in the treatment. Using a single concept shortens both the development time for the treatment and treatment administration time. Using two different concepts, of equal difficulty, on the other hand, permits a partial replication of the experiment relative to the domain of difficult concepts for the age level chosen. In addition, external validity is increased on this dimension as well as conserving experimental subjects in the sense that each subject gives twice as much information as he otherwise would have. If the choice is made to use one concept only, the design remains basically a 2 x 2 x 3 ANOVA. If two concepts are used, it becomes a 2 x 2 x 3 repeated measures, assuming equivalent tests for the two concepts can be constructed.

Following the selection of the concept(s), the subsequent tasks are to vary both the convergence-divergence of the exemplars to be used and the clarity of the characterizations of the concepts. This is done while working within the constraints of a fixed range for the number of moves permitted in each strategy. This range should probably not be fixed exactly at the outset. Rather, the problem of developing examples should be worked on first, then the cut-off points on the number of examples to be given is set according to the difficulty of the concept. It is assumed that mastery of a more difficult concept requires more examples than a less difficult concept. Failure to give enough examples for at least the brighter students to learn the concept well would not provide a completely fair trial of the basic argument, while giving too many examples might obliterate differences between the treatments.

To obtain a range of variation in examples and in the clarity of the characterizations employed, a good initial procedure is to ask students in a methods class in mathematics education as well as expert mathematics educators to write out characterizations and examples and nonexamples of the concepts. This procedure should produce samples of examples and nonexamples and of characterizations. Each sample can then be ordered. The sample of characterizations can be ranked or stratified by clarity. The sample of examples can be ordered on a bipolar continuum from highly convergent to highly divergent, and the sample of nonexamples can be similarly ordered.

To do such ordering, definitions of convergent and divergent examples and nonexamples are needed. These definitions depend both on the series of examples within which a particular instance is buried and on the presence of irrelevant attributes. A convergent positive example is one in which the irrelevant attributes differ only slightly (e.g., by one attribute) from the criterial attribute(s) while a divergent example is one in which many of the irrelevant attributes vary. A convergent nonexample on the other hand is one in which the irrelevant attributes are similar to those of the initial example, but one criterial attribute is removed. A divergent nonexample varies widely from the initial example, in irrelevant attributes, but continues to remove at least one (or possibly more than one) criterial attribute. As a brief and incomplete example, consider teaching the concept of four to a young child. If the initial example is four green pencil marks, a convergent example is four black pencil marks and a divergent example is four red sky scrapers with windows in them. A convergent nonexample is three green pencil marks, while a divergent nonexample is six purple moons with faces.

Once the samples of examples, nonexamples, and characterizations are ordered, the next step is to select from each sample those specific items which might be employed in the experiment. To take this step the exact teaching strategies to be used in the particular experiment must be considered. Certain constraints and certain options are available. One constraint developed by earlier decisions and procedures is that the exemplification moves to be made are to provide interpretable contrasts between different combinations of types of positive and negative examples. This constraint suggests that at least three combinations of examples and nonexamples must be included such that each combination has balanced (equivalent numbers) examples and nonexamples, and

(a) combination 1 uses all convergent examples and nonexamples,
(b) combination 2 uses both convergent and divergent examples and nonexamples in equal numbers, and
(c) combination 3 uses all divergent examples and nonexamples, excluding the first ones given as examples.

Quite clearly, these combinations are not the only ones possible since any ratio of convergent to divergent examples and nonexamples would be permissible. The combinations chosen, however, provide reasonably interpretable contrasts among the full range of possibilities.

A second constraint from previous decisions and procedures is that the combinations of example-nonexample moves chosen must be fit into different patterns which are interpretable as teaching strategies in the Henderson taxonomy. Ideally, these strategies should be drawn at random from that subset of concept teaching strategies for denotative concepts which involves repeatable characterization and exemplification moves. This subset is not presently explicated, however, and the only practical alternative is to draw at least two contrasting strategies from it on an intuitive basis. For present purposes, let us designate one strategy as a C E ∿ E E ∿ E C E ∿ E E ∿ E C strategy in which the connotative or characterization moves are dispersed among counterbalanced example and nonexample moves. This strategy is thus a "dispersal strategy." The second strategy shifts away from dispersion, and blocks or clusters the moves as follows: C C C E E E E ∿ E ∿ E ∿ E ∿ E. Quite clearly, many other specific strategies involving the same elements in different arrangement could easily be generated. The present design does not directly utilize these alternatives, but can be quickly modified to adapt to them as will be noted at a later point.

The remaining step in the design is to contrast "clarity" in at least two levels. A procedure for obtaining differences in the clarity of characterization moves was discussed earlier and initially, at least, less clear and more clear moves can be assigned to the strategies in a balanced design. To see this point more clearly, Figure 3, which arrays all the elements of the design up to this point, should be examined.

The fact that "clarity" as applied to teaching mathematics may involve not only whether the characterization (C) moves are clear but also whether the entire sequence of moves is clear is now open to an empirical test. To make the test, the instructional programs for each cell, a to l in Figure 3, must be written. These programs are then assigned to a panel of judges, for example, a mathematics education class. Each judge then rates each program on a scale of 1-5 for clarity. These ratings are then collected and assigned to the cells in accord with the type of program on which they were made. An ANOVA of the ratings is then performed. If "clarity" is a function of the clarity of the characterization moves, a main effect for B will be present such that the mean ratings for the less clear cells (a-f) will be less than those for the more clear (g-l) cells. If clarity is a function of some other factor such as strategy, or convergence or divergence, or some combination of factors, it will appear in the other main effects (A or C) or in the interactions (e.g., A x C), hopefully in an interpretable way. Notice that this analysis does not reveal anything about the relation of clarity to learning; it tells only whether or not clarity is orthogonal to strategies and convergence-divergence or correlated with one or the other or both, hence confounded with them in this design.

Assigning subjects to treatments. A distinct advantage to the programmed learning approach to studying teaching strategies is that it permits the random assignments of students within intact classrooms

STRATEGIES

B Clarity	C convergence-divergence	Strategy 1 CE ~ EE ~ ECE ~ EE ~ EC	Strategy 2 CCCEEEE ~ E ~ E ~ E ~ E
Less Clarity		a convergent	b convergent
		c convergent-divergent	d convergent-divergent
		e divergent	f divergent
More Clarity		g convergent	h convergent
		i convergent-divergent	j convergent-divergent
		k divergent	l divergent

Figure 3. Structure of design 1.1.

to the treatments. The randomizing procedure is carried out by assembling the total stack of programs to be handed out in a classroom in random order, then passing them to students, who thereby receive their treatments at random.

This procedure solves four important problems. First, it permits the use of a randomized design with high initial internal validity. Second, it permits the experimenter to treat each student as one replication or unit of observation, thus permitting him to use each student as one degree of freedom. If intact classes are randomly assigned to treatments, as an alternative practice, each class is only one degree of freedom. Third, it adapts easily to covariance or stratified designs as methods to increase statistical precision, since the control variable(s) to be used will be available on a student by student basis. Finally, the randomizing procedure used can be expanded up to about 30 different treatments (one student per treatment in an average classroom), thus permitting many different teaching strategies on the same concept (principle or skill) to be tested simultaneously. In short, the procedure lends itself to greatly increasing the sample of strategies used in any one study; hence, it helps in solving the external validity problem associated with the large universe of possible strategies.

Testing learning outcomes. I am going to minimize this aspect of the present design and related designs because an adequate treatment of it requires another complete paper. There are, however, certain aspects of testing outcomes that cannot be overlooked even here.

First, the method and the problems by which learning is to be assessed must be considered, and should be rather fully developed, at the time the concepts to be included in the treatment are selected. For example, if "fourness" were the concept in the present design, the test of the outcomes would in all likelihood be the correct discrimination of fourness among many divergent examples and nonexamples of fourness. The test would be almost certain to closely resemble the treatments delivered in cells *e* and *k* of Figure 3. Such a test might be considered an inadequate or even biased assessment of concept learning. The time to make this judgment is, of course, before running the experiment, not afterward.

Second, if multiple tests (sets of indicators) of the learning outcomes were used, and in the event these indicators were somewhat heterogeneous, i.e., did not describe a unitary factor, the statistical analysis associated with the design would shift from univariate analysis of variance to multivariate analysis of variance.

Third, in the event two concepts were taught in the present design (i.e., a partial replication was done) and were tested with apparently parallel tests, the scores of the tests would have to be equated through a standard score transformation before initiating the analysis. Otherwise, the difference between the concepts, as treated, would be confounded with possible differences between the tests of concept learning.

Expected outcomes. With respect to the argument that teaching strategies affect learning outcomes, the expected outcome for this design is a significant difference between strategies. Additionally, significant effects for levels of clarity and for types of examples-nonexamples would be anticipated. An acceptable and illuminating set of outcomes would be a significant A x C interaction (strategies by types of examples-nonexamples) since this effect would still indicate that the sequence as well as the quality of the moves is important. A theoretically disappointing but empirically interesting outcome would be a main effect for clarity and no other significant results.

Related experiments. Depending on the outcomes of the type of experiment described above, a number of related experiments can easily be generated. If the anticipated outcomes hold, the same experiment can be repeated for different concepts and/or different age groups. If the convergent-divergent examples effects are strong, this line of research can be extended by examining different ratios or different numbers of examples and nonexamples, convergent and divergent. If the strategies effect is strong and the other effects weak, an interesting series of experiments might be conducted to determine the most efficient strategies for attaining a given outcome by manipulating either the number, repetitions, or types of moves made. Whatever series might be chosen, the theoretical objectives are similar -- to verify or to modify and refine the current theoretical structure so that it more sharply explains the relationships between the moves of the teacher and the learning outcomes of students in mathematics.

Design 2.1: Dyadic Definition of "Strategies"

In this design the intent is to carry out research which is as closely analogous to design 1.1 as possible, but employs a dyadic or student response-contingent definition of teaching strategy rather than a monadic definition. Quite clearly, design 2.1 cannot be exactly parallel to 1.1 since the exact sequence of moves, the strategies, cannot be directly controlled. Rather, the strategies must be left as contingencies. Moreover, neither convergence-divergence nor clarity can be controlled since these are also a function of the teacher and left contingent. What can be controlled, albeit imperfectly, is the degree to which the teachers in the experiment have been trained to (a) perform moves relevant to teaching concepts, (b) give convergent and divergent examples and nonexamples, and possibly (c) be clear.

Design structures, cost and intent. The structural features of design 2.1 might be very similar to those of 1.1 if training capabilities and the cost of conducting the experiment were to be completely discounted. In this event, the structure would be that in Figure 4.

This design structure differs from that in 1.1 in that convergent-divergent example and nonexample giving is not divided into three levels, but only two. The reason this factor is collapsed is that training teachers to give all convergent, all divergent, or balanced convergent-divergent examples and nonexamples probably cannot be controlled at the point of application. That is, even though one might try to train teachers to act exactly in accord with these treatments, when the teachers actually teach they might or might not act in accord with their training. The treatments are therefore simplified so that some teachers are trained

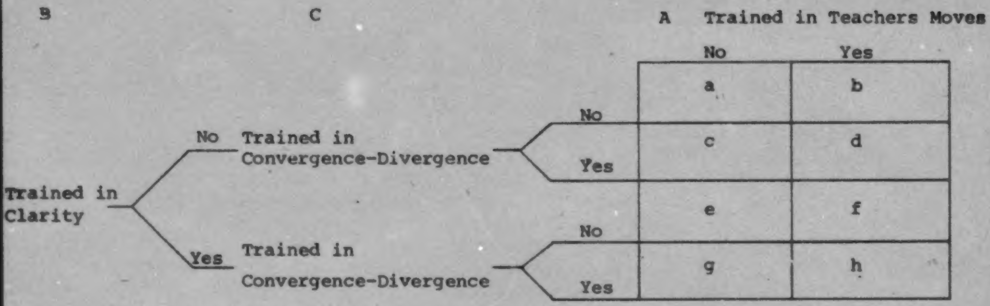

B
C
A Trained in Teachers Moves
No
Yes
a
b
c
d
e
f
g
h
Trained in Clarity
No
Yes
Trained in Convergence-Divergence
Trained in Convergence-Divergence
No
Yes
No
Yes

Figure 4. Possible structure of design 2.1.

in giving convergent and divergent examples and nonexamples and some are not thus trained.

In addition to this structural change between designs 1.1 and 2.1, there are certain other features of 2.1 which are questionable. The first is whether or not one could, within reasonable time limits, successfully train teachers to be "clearer." This seems doubtful, since exactly what one must do to increase his clarity of presentation, questioning, example giving, and the like is poorly understood. To say that one can discriminate between clear and less clear teaching is different than saying that one can train someone to be clear. Working on the latter notion, the experiment could drop clarity as a training treatment and, instead, rate the clarity of teaching actually done in the experiment. In this way, clarity could be treated either as an outcome of the other two factors in the design or as a control variable (covariate) or both.

Dropping clarity as a training treatment would also have a beneficial effect on the costs of the experiment since four of the cells (e, f, g, h) can be eliminated and correspondingly fewer teachers and students included in the experiment. In many experiments, for example, those like 1.1, this cost reduction would be small. But in 2.1 the participating teachers must be both trained and subsequently observed or videotaped and rated. Thus, reducing the number of participants can effect substantial savings which can be diverted to additional experiments.

The second feature of design 2.1 open to question is whether or not training someone in giving convergent and divergent examples and nonexamples is in effect training him to make teaching moves (example and nonexample moves), which of course it is. Nonetheless, a distinction can be maintained in training between teaching someone to vary the quality of his example-nonexample moves and simply teaching him to make example and nonexample moves. If this distinction is not maintained in 2.1, the validity of the design relative to its intent will be compromised. This compromise will occur if in training teachers to give examples and nonexamples, as well as other types of moves, the experimenter inadvertently emphasizes the types of examples or nonexamples to be given.

Let us suppose, however, that the experimenter believes that he cannot avoid addressing the quality of the examples and nonexamples given in his training on teaching moves. If this were to be the case, he might again shift the structure of the design so that only two groups (cells a and b) were involved, thus dropping the convergence-divergence of examples and nonexamples as an explicit factor in the design. In keeping with the intent of both design 1.1 and design 2.1, however, one might reasonably insist that convergence-divergence of examples and nonexamples be taken into account. To meet this intent, the experimenter would observe and rate the teaching of the trainees on the convergence-divergence factor. He would in turn use these ratings in two ways. First, he would test the differences between the mean convergence-divergence of examples and nonexamples given by teachers in cell a versus cell b. This test would tell whether or not his training influenced this factor. Second, he would correlate the ratings for cells a and b pooled to student learning outcomes. This correlation would indicate the degree to which convergence-divergence influenced student learning outcomes. Finally, he might use the rating as a covariate to increase the statistical precision of his design. The latter move would be made, however, only if the correlation between the rating and student learning was substantial -- say greater than .50.

Treatment construction - teacher training. A curious feature of design 2.1 is that its structure depends greatly on how skillful the experimenter thinks he is in constructing training treatments. For present purposes, the writer will assume that the four-cell structure can be carried out if the appropriate training materials are developed.

An important consideration in developing these materials is their cost. If substantial sums of money are available, the materials might include films or videotapes showing teachers modeling the moves that the trainees are to learn. Programmed materials would follow which require the trainee to both discriminate the appropriate behavior in transcripts of teaching and also produce examples of the appropriate behaviors. If modest sums are available, audiotapes or simply transcripts may be used. If virtually no money is to be had, the experimenter may himself have to give the instruction in a highly structured replicable way.

For the treatment which trains teachers to make relevant moves in teaching concepts, the experimenter has the option of teaching all 16 concept moves if he chooses. He may also choose to teach only a few moves, but certainly more than just example and nonexample moves, which would increase the risk of confounding the two treatments in design 2.1. The investigator also has the option of developing materials separately for each individual move or for groups or clusters of moves. If many experiments using different types or combinations of moves are planned, as would be the case in a family of experiments related to design 2.1, much is to be said for having separate materials for each move. Different combinations of moves for different training treatments could thereby be easily produced.

An important aspect of training on concept moves is that the investigator must establish a criterion for trainee proficiency in making each move. This criterion might be to identify and differentiate types of moves as well as produce moves with high accuracy on a paper and pencil test. It might be ability to accurately perform each type of move in a microteaching situation. Which is chosen depends substantially on the time and money available to the investigator.

For the treatment which trains teachers to give convergent and divergent examples and nonexamples, the same type of training materials used to teach concept moves is required. However, instead of noting or even mentioning "moves," these materials focus on the importance of giving good examples and nonexamples in one's teaching. Then they go on to demonstrate different types of examples and nonexamples and when they might be used. Again, a criterion for trainee proficiency in giving examples and nonexamples, convergent and divergent, must be developed by the investigator.

Treatment construction - concept teaching. Once the participants in experiment 2.1, cells b, c and d, have been trained, all participants including those in cell a must be assigned to teach one or more than one concept to children in the designated age range. The concept(s) would be selected in the same way as in design 1.1. In design 2.1, however, all participants must be given access to the same resource material about the concept(s) and given equal opportunity to prepare the instructional lesson. Finally, each teacher must be assigned at random to one or more microteaching sessions, to which the pupils also have been randomly assigned, with the number of sessions determined by the number of concepts used. Following each session, the students are tested for concept acquisition, while the videotapes of the session are saved and subsequently rated for clarity. In addition, the tapes also may be examined or rated relative to the strategies employed and the convergence-divergence of the examples and nonexamples used.

Interpretation of results. The statistical analysis used in design 2.1 might be a two-way analysis of variance, with repeated measures if more than one concept is taught, or analysis of covariance if clarity is significantly correlated with student outcomes. Under any analysis, however, the design should be about equally interpretable. If the main effect for concept moves is significant, then training teachers to make these moves aids student concept learning. Whether this outcome is a consequence of one or more specific strategies being employed by teachers can be determined, however, only by an analysis of the videotapes. The same is true if a significant main effect occurs for the convergence-divergence factor. Perhaps the more likely outcome of design 2.1 is a significant interaction effect which results from highly superior performance by the group (cell *d*) trained both on moves and on convergence-divergence of examples and nonexamples.

Summary

The design of research on teaching strategies requires attention to four domains of variables: teachable objects, student attributes, teaching moves and strategies, and learning outcomes. Within each domain a further classification of variables is required. These classifications should produce sampling frames from which elements may be drawn either to increase the possible generalization of the research findings within the domains or to increase the precision of the experiments to be conducted. The design of experiments involving these sets of variables requires detailed attention to the method of selecting the variables to be employed from each domain, to the procedures for operationalizing these variables, and to the control of variables which potentially offer competition with "teaching strategies" as alternative explanations of the results of the experiments.

References

Cooney, T. J. Rationale for the workshop on research on teaching strategies. (From a proposal submitted to the National Science Foundation for the establishment of the Georgia Center for the Study of Learning and Teaching Mathematics.) Unpublished manuscript, University of Georgia, 1974.

Cooney, T. J., Davis, E. J., & Henderson, K. B. Dynamics of teaching secondary school mathematics. Boston: Houghton Mifflin, 1975.

Dodd, C. A. An exploration of an evaluation model applied to the Indiana University-Mathematics-Methods Program involving the analysis of three dimensions of teaching style and other selected variables and their relationship to pupil achievement (Doctoral dissertation, Indiana University, 1974). Dissertation Abstracts International, 1975, 35, 4610A. (University Microfilms No. 75-1691)

Henderson, K. B. Concepts. In M. F. Rosskopf (Ed.), The teaching of secondary school mathematics. The Thirty-third Yearbook of the National Council of Teachers of Mathematics. Washington, D. C.: The Council, 1970.

Henderson, K. B. The use of pedagogical models. Paper presented to the Special Interest Group in Mathematics Education at a meeting of the American Education Research Association, Chicago, April 1972.

Lester, F. Teaching skills in mathematics instruction. In R. L. Turner (Ed.), General catalog of teaching skills. Albany, N.Y.: Multi-State Consortium on Performance-Based Teacher Education, New York State Department of Education, 1974.

Markle, S. M., & Tieman, P. W. Conceptual learning and instructional design. Journal of Educational Technology, 1971, 1, 1-12.

Nuthall, G. A. An experimental comparison of instructional strategies in the teaching of concepts (Doctoral dissertation, University of Illinois, 1966). Dissertation Abstracts, 1967, 27, 4137A. (University Microfilms No. 67-6694)

Smith, B. O., & Meux, M. O., Coombs, J., Nuthall, G., & Precians, R. A study of the strategies of teaching. Urbana: Bureau of Educational Research, College of Education, University of Illinois, 1967.

Smith, B. O., & Meux, M. O. A study of the logic of teaching. Urbana: University of Illinois Press, 1970.

Thiagarajan, S. Instantial and contrastive teacher moves and student's conceptual behavior (Doctoral dissertation, Indiana University, 1971). Dissertation Abstracts International, 1972, 32, 5662. (University Microfilms No. 72-12,977)

A Research Context for Delivery Systems Research on Strategies for Teaching Mathematics

Kenneth A. Retzer

Illinois State University

A Utopian Fable

The Utopian States has a problem. A team of researchers has, by studying human behavior, identified a regimen of physical exercises which can benefit a substantial portion of the population of Utopia--provided one can trust the preliminary indications of research.

The researchers are enthusiastic; dissemination of these regimens of exercise could be means by which they make a contribution to humanity. Excitement is tempered by honest concern about public acceptance of the potential of these regimens. They know that people are already getting exercise in numerous ways--both structured and unstructured--and they realize that they need to convince some that structured activity is more effective than unstructured exercise without having them feel their freedom or artistic expression is being restricted. They need to convince others who believe in structured exercise that these regimens are more beneficial than those they are presently using.

Furthermore, they are concerned about the point at which there will be enough research evidence to convince them that some of the regimens of exercise might be more beneficial to them and to those they train than those they are presently using. How can they get the financial support and the time for that needed research, and how can they coordinate those research efforts to make the maximum use of funds, mutual communication, and support and, simultaneously, minimize overlapping of efforts and insignificant research?

This fable about the Utopian States has a counterpart in the United States. If exercise can be compared to teaching activity classified as teaching moves and regimens of exercise to teaching strategies which result from sequencing those moves, we may be able to visualize the nature

and the state of the movement in strategies for teaching mathematics. We see, as the Utopian researchers do, the need for adequate research to identify effective strategies and means of sharing the research with preservice and inservice teachers.

> *While the problems of research support and research coordination are formidable ones, they are solvable. As evidence mounts, the researchers will be able to decide whether to continue testing the effectiveness of specific exercise regimens and whether to use this evidence in an effort to convince others to use those regimens identified as effective.*
>
> *Problems of implementation and dissemination arise. Surely the Utopian public schools are the single most promising dissemination source since they are already involved in teaching physical activity. Would not the physical education departments be the most promising departments within the public schools to learn the moves involved and to teach their students to sequence these moves into regimens of exercise? This possibility raises numerous problems and challenges. While not discounting these avenues, the researchers do not feel that dissemination by community organizations, federal agencies, or interested individuals can have quite the wide-spread effect as the schools.*
>
> *Of course, other departments such as home economics, industrial technology, and art are also concerned with physical skills. But one would not expect these departments to be a major source of dissemination unless research makes it evident that the regimens of exercise are most beneficial to those in a particular department. The fact that most of the researchers have backgrounds in industrial technology and that the moves were discovered by observing physical activity related to industrial technology leaves open the question of whether these exercises are primarily beneficial to those in industry as contrasted with the general populace.*
>
> *Among the physical education faculty there are those whose backgrounds are in industrial technology. Some have specialized in exercises beneficial to those in industry. These are among the physical education faculty who would make decisions about implementing the exercises suggested by the researchers. By virtue of their positions, they have an interest in industrially-related exercises as well as feelings about the effectiveness of what they are presently teaching.*
>
> *To our researchers, an analysis of schools and physical education departments is in order. Our researchers are assuming that the physical education departments as well as other departments have curricula which are based on research. Before they would adopt new regimens of exercise as a part of their curricula,*

> *they would want adequate research evidence of their effectiveness. They are quite prepared for the fact that some, because they would prefer to await the results of large-scale research, will not be convinced by pilot projects and small-scale research. Utopian researchers are quite prepared to implement their regimens of exercise in a few schools and design research which tests whether these exercises help the schools achieve their objectives. Assessing the realism of those assumptions and giving preliminary consideration to the timing and structure of research on educational systems is worthy of consideration.*

The analogy can be continued by comparing the public schools to our teacher training institutions, the physical education departments to colleges (or departments) of education, and the industrial technology departments to mathematics (or mathematics education) departments.

By doing so, one can perceive a web of interrelationships between research and dissemination. Researchers, committed to exploring the relative effectiveness of strategies for teaching mathematics, can coordinate a spectrum of small-scale studies. Such studies can point toward large-scale studies to determine whether knowing moves for teaching content increases teacher effectiveness. Simultaneously, there is a need for dissemination of existing classification systems of teaching strategy moves. This dissemination is both a means and an end. It is a means in the sense that it is prerequisite to large-scale validation studies. It is an end in the sense that strategy moves have already exhibited sufficient potential and usefulness to warrant their being shared. The effectiveness of delivery systems used to share strategy moves can be tested. This interrelatedness of research and dissemination efforts warrants further exploration. That exploration is the task of this paper.

The Task

A major purpose of this paper is to explore aspects of research with and dissemination of classification systems of moves for teaching mathematical concepts, generalizations, or skills. Specifically, this paper suggests (a) that validation studies of the type Dossey (see the Dossey paper in this monograph) describes be continued, (b) that similar validation studies in other subject matter fields be considered, (c) that large scale validation studies on strategies for teaching mathematics be anticipated, (d) that dissemination of strategy move classification schemes is a prerequisite for large-scale validation studies, and (e) that delivery system research should be done on the effectiveness of the resultant teacher education programs. This paper is largely directed to discussing aspects of that dissemination and subsequent delivery systems research.

"Validation research" refers to that research which uses sequences of teaching moves in an effort to determine effective strategies. In the context of this paper, "delivery system research" means research on one or more components of a teacher education program to determine its effectiveness.

Exploring aspects of research with and dissemination of classification systems of moves for teaching mathematics is carried out in the following manner. Attention is given to the following:

1. the nature of strategies for teaching mathematics and their source;
2. promising aspects of using those strategies in teaching and in teacher education;
3. some positions and suggestions for a research framework;
4. facets of implementing these strategies in some mathematics teacher education programs including research bases of teacher education curricula, modeling as a dissemination approach, and a case for materials development; and
5. some dimensions of a suggestion for anticipated delivery system and validation research in mathematics education on teaching strategy moves.

Nature and Source of Strategy Moves

A strategy for teaching mathematics is a sequence of bits of verbalization called moves. The strategy moves in current usage were identified by analyzing tapes of classroom discourse with a logical perspective--keeping in mind whether a mathematical concept, principle, skill, or fact was being taught. Henderson offered classification systems of these moves in articles and chapters of professional publications as early as 1967. The most complete single source is a secondary mathematics education textbook by Cooney, Davis, and Henderson (1975). This text contains classification systems of moves for teaching mathematical concepts, mathematical generalizations, and mathematical skills within one volume and reflects the latest thinking of these mathematics educators.

Promising Aspects

Mathematics teachers and mathematics educators have been drawn into the movement for various reasons--including philosophical perspective, aesthetics, applicability, research results, and personal experience.

Some find themselves in agreement with Smith (1971) who stated:

> Almost from ancient times it has been assumed that the way teachers should behave in the classroom could be derived from what philosophers, and in recent decades psychologists, said about thinking and learning. . . .While this approach to the formulation of teaching skills has not been abandoned, it has been challenged as an exclusive approach by research workers who conceive of teaching behavior as worthy of study in its own right. (p. 3)

Mathematics educators who share Smith's perspective are drawn to the results of Henderson's observations of interaction in mathematics classrooms.

Some who have been drawn into the movement find an intrinsic beauty similar to that in the structure of mathematics. Many associated with mathematics appreciate that an endless number of arithmetic facts can be generated using algorithms and less than 100 basic facts. Similarly, some find the fact that an endless number of theorems can be proved with a small finite number of axioms aesthetically pleasing. Likewise, some can appreciate beauty in the determination that the multiplicity of teaching episodes can be interpreted as permutations of a small finite number of strategy moves.

Some are attracted by the applicability of these moves to teaching mathematics content--concepts, generalizations, and skills--regardless of the generalizability of these moves to other subject matter areas. Until the extent of generalizability is known, there is a possibility of finding teaching moves which are unique to mathematics. The opportunity to work with specialized teaching moves which are used to teach mathematics content as contrasted with generic teaching moves which relate to classroom management, structuring materials, and manipulating pupil activities is appealing.

Since Dossey (in this monograph) provides comprehensive details, it suffices to indicate that some are drawn to the classification systems by the encouraging results of some research studies. These results are interesting in contrast to a quotation from Rosenshine and Furst (1971) who stated:

> Educational researchers have not provided those who train teachers with a repertoire of teaching skills which indicate to a teacher that if he increases behavior x and/or decreases behavior y there will be a concomitant change in the cognitive or affective achievement of his students. (p. 40)

Some teacher educators have used the strategy moves to provide a common language to describe teaching and general methods explicitly. Preservice and inservice teachers have reported that knowledge of the moves helps them (a) to write objectives, (b) to plan lessons, (c) to determine alternatives when classroom developments require that they change from their planned strategies, and (d) to write test items. There are adequate subjective reasons for thinking mathematics teachers would find strategies for teaching mathematics aesthetically pleasing and useful. Hence, at least some mathematics educators would want to include them in their teacher education programs.

Positions and Suggestions

A position taken in this paper is that those interested in research on strategies for teaching mathematics should be concerned with the challenges of dissemination and delivery system research on strategy moves from the outset of cooperative research efforts. Delivery system research is as important as validation studies--even prerequisite to large-scale validation studies. The development and use of protocol and training materials in teacher education programs and dissemination of such materials and programs are prerequisite to research on the system's effectiveness and the production of enough teachers to permit large-scale validation studies. The paradigm given in Figure 1 may further explicate this position as well as outline a research framework. In Figure 1, Item (1) designates our present state. Dossey's paper summarizes these validation studies, and his and other papers in this monograph make suggestions for further research identified as (2).

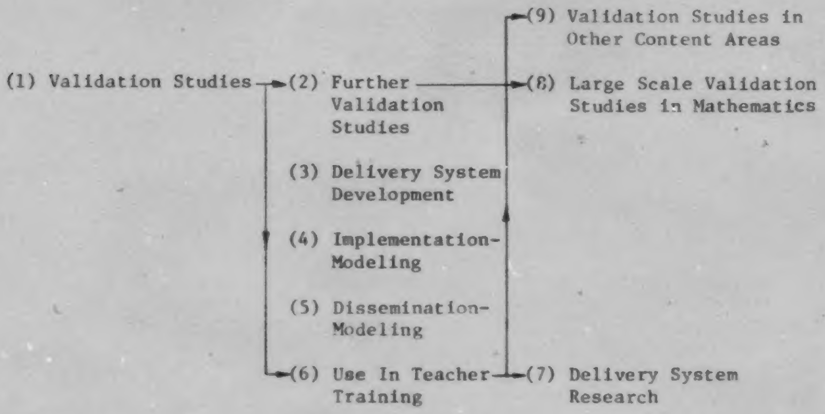

(9) Validation Studies in Other Content Areas
(1) Validation Studies
(2) Further Validation Studies
(8) Large Scale Validation Studies in Mathematics
(3) Delivery System Development
(4) Implementation-Modeling
(5) Dissemination-Modeling
(6) Use In Teacher Training
(7) Delivery System Research

Figure 1

This paper points towards Item (7) and suggests that Items (3), (4), (5) (not necessarily in that order) are important preconditions. Items (3), (4), and (5) could also lead to (6) for teacher education institutions who do not consider (7) or (8) prerequisite to curricular change. It is suggested that (3), (4), (5), and (6) are also preconditions for (8). Decisions need to be made on the value of (9) as desirable research and whether it should be encouraged after (2) or else (8). We need to establish several things to implement teaching strategies into teacher education programs (Item (6) of Figure 1) on a research based rationale. First, are some strategies more effective or more efficient than others (Items (1) and (2))? Second, can teachers be trained to recognize and use moves (Item (7))? Incorporating teaching move classification systems in comprehensive teacher education programs may also require both small (Item (9)) and large-scale validation studies in a number of other subject matter areas.

With this suggested research framework in mind, facets of implementation, dissemination, and prerequisite materials production can be examined. It is asserted that implementing classification systems of moves into some teacher education programs is justifiable on subjective bases. Most components of existing teacher education programs are not based on research evidence. However, the current incompleteness of effectiveness research on knowledge of teaching strategy moves need not be a deterrent to their implementation.

Modeling as a Dissemination Device

It is significant that curricular development in school mathematics content has been accompanied by dissemination efforts. The development of a secondary mathematics curriculum by the University of Illinois Committee on School Mathematics (UICSM) was accompanied by academic year institutes (AYI's) to train teachers (about 1957-1969). The task for these teachers was to become thoroughly acquainted with the UICSM curricular materials and their mathematical bases. Later (circa 1971) the Colorado Schools Computer Science Curriculum Development Project (CSCSCDP) was accompanied by resource personnel workshops (RPW's). These teachers and administrators became familiar with the implementation of computer programming in junior level high school mathematics and the philosophy of the text. More recently, Indiana's Mathematics Methods Program (MMP) has been accompanied by a modeling program (from about 1972) in which visitors can study the materials, learn to use them, see them being used in a model program, and contribute to curricular development and modification.

These three training programs were financially supported by the National Science Foundation and represent distinct programs in the evolution of NSF's thinking concerning teacher training. After several years of retraining experienced mathematics teachers in AYI's, support of the NSF was directed toward RPW's and the improvement of some preservice education programs. The National Science Foundation recognized the futility of continuing to retrain individual teachers, while teacher education institutions were producing teachers who would shortly be in need of retraining. Hence, the NSF began supporting RPW's whose participants were teachers, department chairmen, and teacher education faculty with the expectation that these personnel would further dissemination by training others. Furthermore, support was given to undergraduate institutions for development of their preservice programs so that their graduates would be capable of using the best of curricular materials and teaching methods. Indiana University's MMP contains a recent and a most promising dissemination device because of the nature of the modeling inherent in it. A critical feature of the MMP includes support for teacher educators to visit the model and engage in follow-up communication, as well as the opportunity for the visitors to provide input to the development of the model and, hence, develop a sense of ownership. These characteristics merit consideration.

The MMP program represents recent thinking of the National Science Foundation toward dissemination. It seems reasonable, then, that serious consideration should be given to support of implementation of teaching strategies into teacher education programs which can serve as a model teacher education program to be visited and modified by mathematics educators. Such programs would not be viewed as demonstrating the ideal. Rather, they exemplify a definable program which is in operation. Visitors could help improve the model program as well as adopt or adapt aspects of that model to their own situation.

Stiles and Parker (1969) spoke of teacher education programs being implanted and imitated with minimal evaluation. Undoubtedly there are teacher training programs which will gladly imitate a promising model program. Faculty at such institutions are in search of fresh ideas but often are burdened with responsibilities which prevent them from doing the necessary research and developmental work. Establishment of model teacher education programs and provision of financial support to visit them and to communicate with their personnel would serve these institutions well.

Many teacher education institutions contain faculty who both want to and are expected to be productive along research and developmental lines. Adoption of another institution's program is highly unlikely in these cases; adaption is more likely. It appears that the adaption will be even more likely if faculty members have an opportunity to assist in development to an extent that they acquire a sense of ownership in part of the new curriculum. Girault (1974) stated, "If innovation is to remain alive beyond the introductory developmental stages, it must be 'owned' by the total system it purports to serve" (p. 1). She defined the system members as initiators, legitimatizers, and maintainers. Noting that teacher education programs exist at the interface of the university and the K-12 school unit, she indicated that the innovative system must be defined so as to include representatives of each unit bordering that interface. She pointed out the need for the maintainers to have the opportunity to communicate with and influence the initiators and legitimatizers. The maintainers should share a sense of accessibility, commitment, interest, and mutuality of influence with the initiators and legitimatizers in order for all concerned to feel a sense of ownership in the innovation. Program owners see themselves as shaping the program, as having easy and informal access to other owners, as being convinced that the program is worthy of dissemination, and as seeing their ownership as widely and publicly recognized. Developers of a modeling program such as the one described above might wisely consider Girault's advice and cultivate ownership among all participants.

Caution should be exercised that model teacher education programs not be considered the only dissemination device. Clarke (1971) made the following observation with respect to model teacher education programs funded by United States Office of Education, "None of the programs planned teacher education as something whose beginning and end were in the institution" (p. 127). Likewise, possibilities for implementation of teaching strategies into programs of professional organizations, inservice programs in public schools, state departments of education, and other institutions beyond the university should not be overlooked.

Delivery System Development

What was the case for innovation in school mathematics might equally be valid for mathematics teacher education innovation. The textbook by Cooney, Davis, and Henderson (1975) which contains classification systems of moves used in teaching mathematical concepts, generalizations, and skills should serve textbook needs in mathematics education programs initially. Implementation might be further assisted by development of materials which enhance the preservice teacher's ability to observe these moves in classroom settings and emulate these moves under controlled conditions. Materials designed to develop observational skills (protocol materials) and teaching skills (training materials) might include concise overviews of the classification systems. They could be illustrated with examples, concise lessons in which many of the moves are sequenced into a strategy, and realistic depiction of classroom settings.

There is evidence of the effectiveness of protocol and training materials similar to the ones suggested for development in mathematics education. Brown (1974) has written:

> For several years, those of us who have devoted much time and energy to the creation of protocol and training materials have been operating on the basis of relatively untested assumptions. For example, there is the assumption that protocol materials in fact lead to the acquisition of meaningful, usable concepts. Similarly, there is the assumption that training materials lead to the acquisition of teaching skills which are likewise meaningful and usable. Both of these assumptions now have some supportive data. More centrally, however, there are the assumptions that protocol and training materials lead to differential outcomes and that given a teaching skill for which there is a strong conceptual base that the combined effects of protocol and training materials are additive. (Foreword)

The Winter 1974 issue of Journal of Teacher Education is devoted to a discussion and justification of the place of protocols in teacher education programs.

Delivery system research could be planned almost immediately on the effectiveness of the protocol and training materials whose development is suggested in this paper. This research would be component research and would be classified even more specifically as delivery system research of the type Brown summarized. In fact, Cooney, Kansky, and Retzer (1975) have made a case for development of protocols on strategy moves for teaching mathematics. They identified moves which might be best to initially portray.

A crucial question arises as to how the development of a delivery system can be financed. Regardless of the source of support for the needed development, the resultant research can provide a more rigorous evaluation of the protocols developed than the subjective evaluation commonly used with recent program development projects.

Context of Delivery System Research

Delivery system pilot studies could begin almost simultaneously with the building of model teacher education programs. These pilot studies can provide formative evaluation of the protocol and training materials being developed. They can also help identify research variables, criteria, and other elements of research design used to test the effectiveness of delivery systems. Since this subsequent delivery systems research is one kind of research on teacher education programs, a literature search was made to determine relationships among proposed validation and delivery system research and prior research in teacher education. One outcome of this search is a perspective of the kinds of research on teacher education programs, and an indication how delivery system research and validation studies are related to them. There is an indication that delivery systems research and validation studies should be able to proceed in a friendly environment because the concerns of these kinds of research are similar to that expressed in the broader literature. Finally, the literature provides some challenges and specific things to consider as mathematics educators design their delivery system research and validation studies. These are examined more fully in turn.

The research literature sampled indicates that there are essentially three positions used in evaluation of teacher education programs. One can study the effects of a component of a program or of the entire program. Others take the position that research should evaluate a teacher education program by studying the teachers produced by it. This type of research seems equivalent to studies on teaching effectiveness and teaching evaluation. Thus, one exploring the educational research context of delivery system research cannot help but explore the context of validation studies as well. Figure 2 provides an outline of this perception as well as a basis on which to subsume suggested research in the teaching strategies movement. With respect to Figure 2, a position of this paper is that anticipated delivery system research can be subsumed under either component or comprehensive research. Expected validation research can be viewed as a subset of research on effectiveness of teacher actions, i.e., research on the products of teacher education programs.

Later comprehensive delivery system research on teaching strategies can be undertaken when a suitable design has been established. Finally, with respect to Figure 2, we note that small scale validation studies are elements of the set of studies of teacher effectiveness.

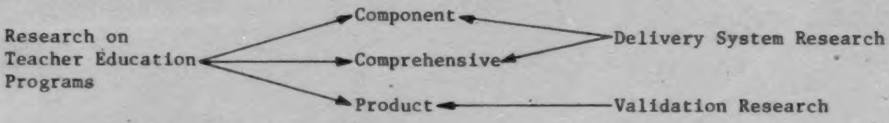


Figure 2

Developing a Research Program on Delivery Systems

Highlights of a proposed sample delivery system will be outlined, and some aspects of pilot delivery research models will be considered. Mathematics educators are challenged to contribute to the development or improvement of this delivery system and subsequent research or else to develop contrasting delivery systems and plan research using them. Next the literature on research on teacher education programs and teacher effectiveness is examined to determine what justification exists for delivery system research and to determine some of the challenges to be faced. Some suggestions are made for evolution of future delivery system models, and some justification for subsequent large-scale validation research is considered.

A Delivery System

It is presumed that all validation and delivery system research coordinated by the Georgia Center will have common reference to the classification systems of moves for teaching mathematical concepts, generalizations or skills which appear in Cooney, Davis, and Henderson (1975). This presumption is made not to inhibit the originality of researchers but to gain the advantages which can accrue from a concentrated and coordinated research effort with a common frame of reference. The sum total of research results from the various projects should contribute substantially to the research literature on teacher effectiveness and teacher education system effectiveness. Such a common frame of reference should also help clearly delineate the nature of the research treatments. The classification systems taught are empirically based; they categorize what mathematics teachers actually do as they verbally deal with mathematical content. This approach clearly contrasts with classification systems of teacher actions based on a philosophy of education or a psychology of learning. It also contrasts with teacher actions which are nonverbal or are related to interpersonal relationships or classroom management functions of teachers. Hence, it is suggested that a delivery system contain these classification systems of moves together with the protocol and training materials necessary to enable preservice or inservice teachers to attain the ability to observe and use these moves.

A suggested perspective for this sample delivery system is that teachers should be trained to be observers of classroom interaction as well as directors of the verbal manipulation of mathematics content. Observing and using strategy moves are pedagogical skills. Hence, they should be taught using a delivery system which uses such skill strategy moves as interpretation, demonstration, and guided practice with feedback.

Protocol materials are intended to develop observational skills. Such materials can help one (a) develop the ability to discriminate the teaching of mathematical content from other aspects of classroom interaction; (b) distinguish among the teaching of mathematical concepts, generalizations, or skills; and (c) identify sequences of moves. It is suggested that the following protocols be prepared for use in developing observational skills:

1. a protocol which gives examples and nonexamples of teachers dealing verbally with mathematical content and which establishes that a teacher, a student, a text, an activity card, or a multi-media presentation can each convey subject matter and, hence, make teaching moves;

2. a protocol which identifies the teaching of mathematical concepts, generalizations, or else skills as contrasted with other types of mathematical subject matter such as facts or values;

3. protocols which depict the major categories of moves for teaching mathematics--such protocols should be prepared for each content category--concepts, generalizations, and skills;

4. protocols providing an illustrated overview of each move in each of the strategy move classification systems;

5. a number of protocols which are designed to provide enough examples of individual clusters of strategy moves that one might expect the user to attain proficiency at identifying the moves exemplified;

6. a sufficient number of concise mathematical lessons depicting the sequencing of as many moves as possible from each classification system into appropriate teaching strategies to exemplify each move in the three classification systems;

7. a sufficient number of protocols depicting realistic sequencing and pacing of strategy moves to exemplify each move in the three classification systems (Such protocols could be used in practice sessions as alternatives to live classroom observation. They would depict content strategies which are sequential, nested, and overlapping, and which are intermingled with such things as teacher managerial or disciplinary functions.); and

8. protocols, similar to ones identified in 5, 6, and 7, to be used in evaluating the extent to which teachers have developed observational skills.

The preservice or inservice teacher could use the above protocols together with supportive printed matter in order to learn the classification systems and receive guided practice with feedback on the skill of identifying strategy moves.

In using the delivery system with preservice teachers, emphasis might be placed on attaining observational skills related to monitoring their own teaching and identifying strategies used by authors of textual and resource materials (as these activities relate to the selection of texts or activity cards). One might also stress using moves in basic lesson planning, microteaching those lessons, and writing test items. On the other hand, experienced teachers might be asked to sharpen their teaching skills by guided practice in adapting lesson plans when the classroom situation calls for it and in communicating with other teachers' suggestions for teaching alternatives and for improvement of instruction. For both preservice and inservice teachers, one could emphasize the observational skills that would enable them to monitor the content of their own lessons as well as provide a cognitive description of the content manipulation in the class of a colleague or student teacher (as contrasted to the more typical affective or evaluative descriptions which result from observation).

Formative evaluation needs to be done during the development period to help determine the proper combination of multi-media used in making the protocols. Observing and using strategy moves are pedagogical skills, and it is unlikely that the printed media alone can best portray these teaching moves. In some cases where verbal interaction is the primary focus, audiotapes may be adequate. In recent years, protocol materials have been developed in slide-sound, video-tape, 8mm movie, and 16mm movie media. Interested researchers should acquaint themselves with the advantages and disadvantages of various multi-media formats and develop the best possible protocols consistent with the level of financial support available.

The delivery system described could be used with preservice teachers in the mathematics methods course which is common to all secondary teacher education programs or else with inservice teachers in short courses, workshops, or extension courses.

Smith (1971) characterized a teacher education program as follows:

> Specifically then, any adequate program of teacher education provides for (1) training in skills, (2) teaching of pedagogical concepts and principles, (3) developing relevant attitudes and (4) teaching the various subject matters of instruction. (p. 2)

The delivery system described would be an integral part of training in skills and would be used in conjunction with the other components listed. Smith continued:

> A major breakthrough in the training of teachers occurred when teaching behavior was conceived to be a complex of skills that could be identified and practiced systematically under specifiable conditions. (McDonald and Allen, 1967) This conception probably arose from advancements in technological devices for recording and reproducing behavior. . . .Along with this breakthrough came a new emphasis in the analysis of teaching

> behavior. . . .These two points of view--that teaching behavior should be analyzed in terms of the psychology of learning and that it should be studied in its own right--are not contradictory, as is often claimed, but complementary. (pp. 2-4)

The perspective that teaching behavior should be analyzed in its own right gave rise to various classification systems of moves.

Delivery System Pilot Studies

Pilot studies on the effectiveness of delivery systems are needed. Work in teaching strategies is sufficiently new that cooperating researchers need time and experience to identify criterion variables, gain experience with alternative treatments, and wrestle with philosophical considerations of appropriate research designs. Pilot studies could be done with preservice teachers and simply test to see if they have achieved desired levels of ability to observe and use strategy moves. One option in such a study would be to follow these teachers through initial years in the profession to see if use of these skills is apparent. Pilot studies could also be conducted with high school supervisors of student teachers to see if strategy moves are used in observing and communicating with student teachers. Pilot studies need not be restricted to teacher education institutions. Mathematics supervisors working for state departments of education could use a delivery system with department chairman or other administrators. An effort also could be made to determine if strategy moves are helpful in evaluating teachers.

Many teacher training institutions are moving toward competency-based teacher education programs. However, most professional teacher organizations have consistently opposed competency-based certification, performance-based reward systems, and accountability requirements. This opposition seems to be based upon a mistrust of the qualifications of those that judge and upon the common knowledge that many factors other than teacher performance affect achievement of the learner. Delivery system pilot studies might indicate whether it is feasible to consider offering a counterproposal that evaluators be required to reach a specified level of observational skill on those competencies agreed to by teachers. This would be an alternative to being held accountable only in terms of pupil achievement.

Several challenges among cooperating researchers are readily apparent. The construction of delivery systems is one. Another is adapting a given delivery system to one's own need. It would seem that cooperation in developing a delivery system and modeling it in varied contexts represents a worthwhile challenge. In addition, delivery systems research designs must be established. Both development and subsequent research are necessary to maximize information and to warrant subsequent large-scale validation studies on the effectiveness of various strategies.

Although it would be desirable to have cooperating researchers work with similar delivery systems for the training of secondary mathematics teachers, adaptations might be easily made for delivery systems research with mathematics teachers at other levels. Similar delivery systems could be used with preservice or inservice community college teachers as a part of Doctor of Arts or Ph.D. programs. The effectiveness of the delivery system in this context could be tested further. Adaptation of the delivery system for elementary school mathematics teachers might be considered. Since the elementary school teacher has less opportunity than the secondary teacher to become familiar with types of mathematical content, delivery systems for elementary teachers might help them develop proficiency in observing and using the major categories of moves in the classification systems.

Results of a Literature Search

Rosenshine and Furst (1971) have suggested that development and research on delivery systems contribute to the study of teacher effectiveness. The work of Cooney, Davis, and Henderson (1975) now makes such study possible. Furthermore, their classification systems enable us to define strategies using specific denotable behaviors--a condition that had not been met when Rosenshine and Furst made this suggestion. Rosenshine and Furst further suggested:

> Perhaps the next step in increasing control in process-product studies would be to stabilize the teacher's behavior through training so that the observed behavior would be a more accurate reflection of the teacher's intentions and/or the intentions of those who prepare the instructional material. Curriculum developers and teacher educators would have to work together on this problem. Without such cooperative work we may continue to have curriculum experts developing instructional packages without clearly specifying teacher behaviors and teacher educators training teachers in teaching skills without clearly specifying the instructional situations in which they will be used. (p. 62)

Using a delivery system of the kind outlined should certainly "stabilize the teacher's behavior through training." The interrelatedness of teacher actions appropriate to teaching specific mathematical content should permit us to do research in which both the content to be taught and the expected teacher behaviors are clearly specified.

An indication of the potential value of the subsequent large-scale validation studies is that they may provide information which Turner (1971) indicated is needed.

> A recent experiment by Nuthall (1968) indicated that although a particular maneuver or strategy of teaching may facilitate the learning of one particular concept, it may not facilitate the learning of another. Thus the consequences of particular maneuvers are at present hard to predict since they interact with the substantive topic. Nonetheless, the way in which the teacher maneuvers within a particular topical venture may be regarded as potentially criterial for the assessment of teacher performance, although substantial work remains to be done to specify relationships between particular kinds of maneuvers and particular types of ventures in specific subject areas. (p. 23)

Research literature reflects an increasing disenchantment with finding an association between teacher performance and personal/social factors such as those measured by the Ryans' Teacher Characteristic Schedule or the Minnesota Teacher Attitudes Inventory (MTAI). Indications of this have been given by Turner (1971) and can be seen in the following quote from McNeil and Popham (1973):

> The single most important deficiency in research on teaching effectiveness is the failure to use outcome measures as a criterion and, instead, to rely upon a priori measures of a teacher's personal attributes such as his personality or education, his background, or the measures of instructional processes such as his instructional strategies or his verbal behavior in the classroom. When one considers the idiosyncratic background of teachers and pupils, the great range in typical instructional objectives, and the immense variation in the environments where teaching occurs it is unlikely that any processes or personal attributes on the part of teachers will invariably produce pupil growth . . .Systems for guiding the observations of teachers and pupils interacting are legitimate tools for obtaining a more accurate account of what is taking place during the teaching act. (pp. 220-221)

The classification systems in the delivery system described earlier should enable researchers to obtain an accurate account of the teaching act for research purposes.

Thus, there are preliminary indications that an evolved delivery system research design can provide a basis for improved educational research. It provides a framework for the study of teachers' actions (as contrasted with attributes). Further, it contains a clearly specifiable and observable set of moves from which one can sequence moves in order to study the relative effectiveness of strategies which are bound to mathematical content. With this in mind, it might be profitable for cooperating researchers to further examine the historical research context which delivery systems research and subsequent large-scale validation studies would build upon and contribute to.

Challenges in Structuring Delivery System Research

Among the considerations which are crucial to evolution of research designs for delivery systems research and subsequent large-scale validation studies are (a) the "can do/will do" phenomenon which appears in teaching, (b) differences in perspectives on what constitutes teaching behavior, (c) desirability of student learning as a criterion variable, and (d) the philosophical quandry about the relationship between teaching and learning. These are examined in that order.

An initial problem confronting researchers will be to show that a preservice teacher can perform teaching moves in a sensible sequence. Another type of research would be to incorporate treatment as a part of a total preservice teacher education program and determine if teachers will use strategies as defined by teaching moves. One could use as criterion variables preservice microteaching performance, teaching acts of the inservice teacher, or else something related to student achievement.

The search for criterion variables is a challenging and serious one. A design in which preservice teachers are expected to demonstrate that they can perform certain strategy moves would utilize a concise and manageable criterion variable. Studies using such a criterion variable should be done to supply an existence proof that one can train teachers to use teaching moves.

There is a nagging concern about the value of such a criterion variable. If such research were proposed in K-12 mathematics rather than in teacher education in mathematics, the corresponding treatment would be teaching a mathematical skill, and the corresponding criterion would be a demonstration of student achievement. Because so many mathematics education experiments on elementary and secondary curriculum during the development of "new math" curricular materials provided evidence that students can learn various concepts, principles and skills of mathematics at various levels, Bruner was inclined to enunciate an axiom paraphrased as, "Any subject matter can be learned at any level provided it is put in appropriate form." To what extent this presupposition can be generalized to teaching skills at the college level with preservice teachers is a question the potential researcher will want to wrestle.

If, on the other hand, the criterion measures are related to observing a trained teacher within a few years after he has received his training, we have implications of both positive and negative results to weigh.

Positive results would indicate that a teacher who was taught moves within strategies for teaching mathematics concepts, principles, or skills not only can, but will, use them in teaching mathematics. One would have to structure his design so as to tell, if possible, whether this teaching performance was because of training in using strategy moves or in spite of such training. In the latter case, the use of the moves may be the result of unconsciously doing what seems to be appropriate--the kind of appropriate teaching which produced the classification systems in the first place.

If after a few years, teachers were not using strategy moves, however, additional questions would be raised. It is a well-known fact in human behavior that what people will do in an unstructured free choice situation is considerably less than what they can do in a structured situation where certain kinds of performance are expected of them. Without adequate data to predict a "usual" can do/will do loss, inadequate performance of teaching moves would be difficult to attribute to inadequate teacher effort or to an inadequate training program.

The question of a conceptualization of teaching is also a challenging one because it relates to treatment variables, Smith (1971) identified a perspective for research on teacher education and related controversies.

> What is research on teacher education? In a sense this question is naive, for everyone must know already what teacher education is and that research on it is simply the systematic study of problems that arise in the course of carrying it on. Generally speaking, research on teacher education attempts to answer the question of how the behavior of an individual in preparation for teaching can be made to conform to acceptable patterns. . . .
>
> The various conceptions of teaching have given rise to theoretical controversies, which in turn pose the question of how teaching is to be conceptualized for research purposes. . . Despite all our efforts, we apparently have no generally accepted conceptual system, psychology or otherwise, by which either to formulate or to identify the skills of teaching. Fortunately, the lack of such a system does not preclude research. . . . On the other hand it is clear that research would be advanced measurably by a conceptual system for formulating and identifying teaching skill. (pp. 2-4)

Several authors might be cited which relate to Smith's perspective. Teaching behavior is a central focus in some model teacher education programs. Johnson, Shearron, and Stauffer (1968) made the following observations regarding this focus.

> Gem's [Georgia Education Model] position is that the teacher education program should be designed in relation to the job the teacher is required to perform in the classroom. By

> defining what the job actually is, the competencies necessary to perform specific tasks may be adequately determined. In other words, it would logically follow that the content of a teacher education program should be based on the teaching act itself. (p. 5)

There are contrasting opinions on how teaching should be characterized for research purpose. Green (1971) used the method of analytic philosophy to conceptualize teaching. He stated:

> In order to find out what teaching is by observing someone doing it, we need to know what teaching is already. . . . In analyzing a practical activity like teaching, the aim is not to invent some new concept or idea of teaching, nor even to specify what people ought to mean by "teaching." The objective is rather to study, clarify, and more thoroughly understand the idea of teaching that we already have. (p. 3)

Johnson, Rhodes, and Rumery (1975) noted that no acceptable conceptual system exists for characterizing teaching:

> The current approaches to the evaluation of teaching can be grouped in three broad categories: (1) measurement of learning outcomes presumed to be the result of teaching: (2) measurement of teacher characteristics presumed to facilitate learning or the attainment of other possible educational goals; and (3) analysis and measurement of relevant categories of pedagogical behavior. In the sections of this essay that follow, we will attempt to show that these three approaches to the evaluation of teaching have "reached a dead end," not because they have been technically misapplied but because they are fundamentally misdirected. However, all three of these approaches also share two basic problems: (1) the absence of adequate theoretical development or integration and (2) the confusion of measurement with evaluation. (p. 176-177)

It is this "absence of adequate theoretical development or integration" that offers mathematics educators an opportunity to contribute to a theory of teaching which can clearly enunciate a position on its conceptualization.

A third consideration relates to the desirability of student learning or student gains as a criterion variable. Peck and Tucker (1973) wrote:

> One long-needed methodological advance is beginning to appear in research: the use of pupil-gain measures as the ultimate criteria of the effectiveness of any given process in teacher

> education. These include affective and behavioral gains as well as gains in subject mastery. (p. 943)

Rosenshine and Furst (1971) used "process-product studies" to describe research whose criterion is what Peck and Tucker described as "long-needed methodological advance. . .pupil-gain measures." They cautioned:

> The results of process-product studies must be treated with caution because these are correlational, not experimental, studies. The results of such studies can be deceptive in that they suggest causation although the teacher behaviors which are related to student achievement may be only minor indicators of a complex of behaviors that we have not yet identified. (p. 42)

Yet Rosenshine and Furst seemed to encourage a continued investigation of the relationship between teacher behaviors and consequent student learning in the following passage:

> The descriptive behavioral data obtained from these classrooms studies is then compared with what educators believe "should" occur in classrooms. Teacher training then becomes a procedure for closing the gap between the behaviors which do occur and the behaviors which educators believe should occur by training teachers in the desired behaviors. . .Unfortunately, the relationship between the teacher behaviors advocated by educational experts and the consequent learning by students has not been thoroughly investigated. (p. 39)

McNeil and Popham (1973) gave some indication of the widespread acceptability of student learning as a criterion variable by stating:

> A focus on pupils reveals far more about the effectiveness of teachers than does direct study of teachers themselves. . . . Support for the position that the ultimate criterion of a teacher's competence is his impact upon the learner has been offered by a number of individual researchers as well as professional associations. (p. 218)

But McNeil and Popham (1973) also gave us a hint of the difficulties in accepting pupil learning as the chief criterion of teacher effectiveness.

> But reservations in accepting pupil change as the chief criteria of teacher effectiveness have arisen both from technical problems in assessing learner growth and

> from philosophic considerations. Chief among the former are concerns about the adequacy of measures for assessing a wide range of pupil attitudes and achievement at different educational levels and in diverse subject-matter areas, . . .failure to account for instructional variables that the teacher does not control, . . .and unrealiability in the results of teacher behavior, that is, inconsistent progress of pupils under the same teacher. . . .Philosophic differences, of course, underlie questions about the selection of desirable changes to be sought in learners. (p. 220)

Let us examine one such philosophic difference. Johnson, Rhodes, and Rumery (1975) asserted that measurement is confused with evaluation. These authors indicated that the use of the term "evaluation" seems to imply the requirement of a normative theory since value judgments are necessarily involved. They pointed out that the common use of the word "measurement" is essentially descriptive, not formative. Thus, a crucial aspect of evaluation of teaching is the establishment of the relative worths of alternative outcomes. Researchers should make the value system upon which they base teacher evaluation explicit.

Johnson et al. (1975) also examined and refected several possible connections between teaching and learning.

> The logical basis for the use of measures of student attainment as either proximate or ultimate criteria of teacher effectiveness seems to be represented by the hoary slogan if the student has not learned, the teacher has not taught or as is sometimes succinctly stated, "no learning, no teaching." (p. 179)

They claimed that the logical basis for the criterion of student attainment is the contrapositive of "Teaching implies learning." This would be a false implication if we will admit to a situation where teaching does take place with learning not taking place. We would want to reject teaching as a sufficient condition for learning. Elementary logic texts warn students not to identify "implies" with "causes." While "Teaching causes learning." is a relationship to be considered, we would probably want to reject it on the same basis for which we reject "Teaching implies learning."

Johnson et al. (1975) seemed to make the logical error of reconstructing "If a student has not learned, the teacher has not taught." as "Teaching has occurred if and only if learning has occurred." Apparently, they then rejected the resulting equivalence with a plausible example of learning without teaching (which, incidently, would better falsify the proposition "Learning implies teaching" but which adequately falsifies the equivalence).

Johnson et al. (1975) characterized Dewey's position in How We Think (1910) as suggesting that teaching and learning are correlative and comparable to buying and selling. But they insightfully pointed out that nothing bought has not been sold. It is unreasonable to assert that nothing learned has not been taught. They admitted that:

> There is an admittedly plausible sense of which if we do not intend that students learn as an ultimate consequence of what we call teaching activity, then to engage in teaching would be odd behavior indeed. But buying and selling have a similarly loose sense. It is plausible to say that in the market there are buyers and sellers who cannot always buy or sell even though this is what they intend. Both teaching-learning and selling-buying lack precisely that necessary connection which they must have if we are to make warranted inferences from one to the other as evaluation by outcome purports to do. (p. 181)

Another possible connection might be on statistical rather than logical grounds. According to Johnson et al., "teaching may be considered effective to the extent that it increases the probability of specified learning outcomes" (p. 182). But this belief must be subjected to the test of showing that other alternatives to the observed learning are less plausible or probable, e.g., cheating.

Enough literature has been examined to question the criterion of student learning. It is hoped teaching strategy researchers will give some attention to this question also.

Future Directions for Delivery System Research

Once philosophical and design questions have been settled and effective strategies are identified, observational schemes could be enlarged to include aspects of classroom atmosphere and interaction other than content manipulation. Current studies have focused only on content manipulation as defined by teaching moves. Rosenshine and Furst (1971) pointed out that current observation instruments do not record the context of the teaching act, for example, assignments students wrote or the physical environment of the room. They suggested that observation instruments be modified to record a distinction between a teacher's academic directions and disciplinary directions and a distinction between a simple repitition of a student's statement and a teacher's summary of it. Furthermore, they suggested it may be interesting to study time or frequency as the analytic unit. Rosenshine and Furst (1971) also suggested the desirability of having instructional periods ranging from 15 minutes to 10 one-hour daily lessons. Such periods may allow researchers to better focus on specific aspects of the teaching act such as explicating new material or contending with classroom management problems.

Eventually, research on entire teacher education programs should be conducted. Clarke (1971), in a review of research on teacher education programs, noted the immense complexity of variables in such research. Some of these variables (referred to as presage variables) include decisions which precede the design of programs, the contexts of programs, cybernation or self-corrective devices, or the question of who controls the program, the nature and extent of the boundaries of teacher education, and selection procedures. In addition to these presage factors, Clarke also identified product factors. Clarke made the following observation with respect to product factors as they relate to research in teacher education programs.

> The product factors, or teacher behaviors to be produced, were specified in many of the sources reviewed. Designs to evaluate these behaviors were not, on the whole, well-developed, with the exception of one model teacher education program. Evaluation and feedback on the process of teacher education changes still need to be made in the individual candidate's behavior, while evaluation and feedback on the product call for corresponding activities in the light of the candidate's performance in the field. (p. 153)

In summary Clarke noted that

> the prospects for research on teacher education programs are bright. The Models have provided considerable development of theory. Modules lend themselves to micro research. There is a ferment of activity in teacher education. The most difficult area is research on the total program in terms of success in the field. It is suspected that studies of this nature often remain unpublished. (p. 154)

The lack of comprehensiveness in designs of the models was pointed out by Clarke as follows:

> The common complaint is that the content and treatment are too frequently designed to prepare the student for further study of the disciplines. (p. 125)

> Considerably less than half of the designs are proposals for the preparation of teachers reviewed includes serious consideration of the integration of the general education, subject matter and related discipline components into a total program of teacher education. (p. 127)

Finally, we return to the thought that a desirable outgrowth of delivery systems research is the training of an adequate number of teachers for large-scale validation studies. These kinds of studies have justification in recent research literature.

Rosenshine and Furst (1971) pointed out weaknesses in some of the more recent subject matter studies and made a suggestion.

> We need studies in which (1) the teacher is the statistical unit of analysis; (2) teachers or classes are randomly asigned to treatment; (3) observational data are obtained on the fidelity of teacher behavior to experimental or contrast treatment and on the behavior of the student, while similar observational data are obtained on events in the classrooms of teachers who follow the normal procedures; and (4) student performance is assessed by a variety of end-of-course tests. Such studies are rare. To date, we have found no more than ten studies which satisfy all four criterion. The scarcity of such studies is not surprising because conducting them involves enormous problems of methodology, administration and teacher training. (pp. 41-42)

Their suggestion helps justify the goals of large-scale validation studies and comprehensive teacher education delivery system studies suggested in this paper.

Summary

The structure and major points of this paper could be summarized as follows. An attempt was made with the fable of the Utopia States to establish, by analogy, a perspective of the strategies for teaching mathematics movement--both its concern and its current status. Agreement on concerns and status is a prerequisite for future cooperative efforts in research.

At appropriate stages, research needs to be of two types: validation studies and delivery systems studies. The validation studies are research on the effectiveness of strategies which are described in terms of the Cooney, Davis, and Henderson (1975) classification systems of moves. A decision needs to be made as to when a sufficient number of validation studies have been done to warrant moving toward delivery system research. Delivery system research would incorporate strategy moves in a mathematics teacher education program and do research on the effectiveness of the resultant teacher education program. Prerequisites to delivery system research are the implementation of strategy moves into teacher education curricula and the development of components of that delivery system which can facilitate this implementation. The literature on teacher education programs indicates that implementation of teaching strategy moves need not wait until a complete set of validation studies are completed. Current teacher education programs generally are based on subjective judgments of teacher educators rather than on components which have been shown effective by research. Modeling is suggested as a dissemination device. Implementation, program development, and modeling are seen as prerequisites not only to research in teacher education but larger validation studies where research populations can be teachers rather than students.

An attempt was made to present a picture of the larger context into which validation and delivery systems research would fit by sampling pertinent literature on the effectiveness of and evaluation of teaching as well as research on teacher education programs. A sample system was described , and the nature of pilot studies on its effectiveness was explored. Good teacher education studies are relatively recent delivery system research on strategies for teaching mathematics should be a welcome addition to that research--if one can judge by the concerns and suggestions of reviewers of current research.

Challenges related to research design were examined. One type of teacher education research attempted to determine if teachers can perform as they are instructed. The question was raised concerning the value of similar research with respect to performance of teaching strategy moves and concerning research which will demonstrate what teachers will do as contrasted with what they can do. The absence of an agreed conceptual framework on the nature of teaching upon which to base research was noted. The promise as well as the reservations about pupil gain measures as criterion variables of teacher effectiveness were examined. The question of the relation between teaching and learning was raised because of the attention the literature has given to learning as a criterion variable on validation studies.

Discussion Questions

To help us determine (a) the accuracy of the perceptions expressed, (b) the desirability of the goals outlined, (c) the validity of the suggested sequential order of the tasks in a cooperative effort, and (d) the adequacy of coverage of challenges in developing appropriate research designs, the following discussion questions are offered. The hope is that discussion of these points can help the coordinated efforts necessary to forward the movement on strategies for teaching mathematics.

1. Does the Utopian fable represent an accurate analogy of the status and concerns of the researchers in the strategies for teaching mathematics movement?

2. What kind and amount of evidence from validation studies would adequately justify structuring teaching by explicitly using strategy moves?

3. Is delivery systems research needed concurrently with additional validation studies?

4. Are implementation efforts necessary prerequisites to desired delivery systems research?

5. Does establishment of model teacher education programs offer the most promise for dissemination efforts?

6. Is development of protocol and training materials a necessary or desirable part of implementation and dissemination efforts?

7. Would studies establishing whether preservice teachers can use strategy moves be a valuable type of delivery systems research?

8. At what stage should delivery systems research be designed which observes what teachers trained in using strategy moves actually do in the mathematics classroom and the resultant effects upon their students?

9. What should be the role of researchers on teaching strategies in development of protocol and training materials and in persuasive writing which could be used in dissemination efforts?

10. Should the strategies for teaching mathematics components of teacher education programs be the only ones tested in delivery systems research? Should research be designed which tests a comprehensive teacher education program taking into account such things as mathematics courses, education courses, and influences of other agencies such as professional organizations, state offices of education, and certification boards?

11. Which of the following should dissemination efforts be directed toward: teacher education institutions, departments of mathematics, departments of mathematics education, departments of education, offices of state superintendents of public instruction, teacher certification boards, or professional organizations in mathematics education?

12. What is the relationship between teaching and learning, and how can anticipated research contribute to knowledge of that relationship without falling into possible pitfalls of prior research?

Discussion of these points may help us in organizing and coordinating our research efforts.

Let us begin the immediate tasks of developing a teaching strategies delivery system, implementing it in several teacher training institutions while simultaneously doing pilot studies on its effectiveness. Only then can we supply enough teachers to do the needed large-scale validation studies and subsequently produce adequate delivery research designs. Let us accept a challenge to cooperate in these tasks which can contribute to a theory of teaching and to establish a firm basis for the education of mathematics teachers.

References

Brown, L. In the Foreword of *Effects Protocol and Training Materials*, by J. Kleucker. Report No. 6. Bloomington, Indiana: National Center for the Development of Training Materials in Teacher Education, 1974.

Clarke, S. C. T. Designs for programs of teacher education. In B. O. Smith (Ed.), *Research in teacher education: A symposium*. Englewood Cliffs, New Jersey: Prentice-Hall, 1971.

Cooney, T., Davis, E., & Henderson, K. *Dynamics of teaching secondary school mathematics*. Boston: Houghton-Mifflin Company, 1975.

Cooney, T., Kansky, R., & Retzer, K. *Protocol material in mathematics education: Selction of concepts*. Bloomington, Indiana: National Center for the Development of Training Materials in Teacher Education, 1975.

Dewey, J. *How we think*. New York: Regnery, 1971.

Girault, E. *Supervising change in preserivce teacher education or: Who owns the action*? Paper presented at the NSTA/NSF Symposium, Washington, D. C., 1974.

Green, T. F. *The activities of teaching*. New York: McGraw-Hill, 1971.

Henderson, K. B. A model for teaching mathematical concepts. *The Mathematics Teacher*, 1967, *50*, 573-577.

Johnson, C. E., Shearron, G. F., & Stauffer, A. J. *Georgia Educational Model Specifications for the Preparation of Elementary Teachers*. Washington, D. C.: U.S.O.E. Bureau of Research, U. S. Government Printing Office, 1968. (FS 5.258:38019).

Johnson, H. C., Rhodes, D. M., & Rumery, R. E. Part I--The Dunkard search: A critical retrospect. In J. A. Gaston & J. R. Kolb (Eds.), *The assessment of teaching in higher education (A critical retrospect and a proposal)* (Vol. 4). Holland: Elsevier, 1975.

McDonald, F. J., & Allen, D. W. *Training effects of feedback and modeling procedures on teaching performance*. Palo Alto, California: Stanford University Press, 1967.

McNeil, J. B., & Popham, W. J. The assessment of teacher competence. In R. M. W. Travers (Ed.), *Second handbook of research on teaching*. Chicago: Rand McNally, 1973.

Nuthall, Graham. An experimental comparison of alternative strategies for teaching concepts. American Educational Research Journal, 1963, 5, 561-584.

Peck, R. F., & Tucker, J. A. Research on teacher education. In R. M. W. Travers (Ed.), Second handbook of research on teaching. Chicago: Rand McNally, 1973.

Rosenshine, B., & Furst, N. F. Research on teacher performance criteria. In B. O. Smith (Ed.), Research in teacher education: A symposium. Englewood Cliffs, New Jersey: Prentice-Hall, 1971.

Smith, B. O. (Ed.). Research in teacher education: A symposium. Englewood Cliffs, New Jersey: Prentice-Hall, 1971.

Stiles, L. J., & Parker, R. P. Teacher Education Programs. In R. Ebel (Ed.), Encyclopedia of educational research. New York: Macmillan, 1969.

Turner, R. L. Conceptual foundations of research in teacher education. In B. O. Smith (Ed.), Research in teacher education: A symposium. Englewood Cliffs, New Jersey: Prentice-Hall, 1971.

The Materials of Teacher Training

David Gliessman

Indiana University

In the vocabulary of this monograph, the task of this paper is to address the question of how to implement effective procedures for training mathematics teachers to use specified teaching moves and strategies in teacher education programs. Stating the problem in this way places it squarely within what Retzer, in this monograph, refers to as "the development of a delivery system." This is a particularly good point of departure because it is also a critical juncture: Successful implementation is partly a test of the foresight that has been shown in developing *usable* training procedures; if such foresight has been lacking, and implementation is unsuccessful, the most *effective* training procedures are inconsequential.

To begin with what is probably an unnecessary word of caution, there clearly are difficulties in implementing new procedures within ongoing programs. In teacher education, introducing and successfully instituting change is a matter about which some have hypothesized and many have despaired. As an example of just how intractable some feel the problem to be, an academic division in the author's own institution has even established a concentration in "Diffusion and Adoption of Innovations." The implication of that kind of action is clear; mathematics educators will not be the first nor the last to try to bend, mutilate, or fold the "delivery systems" that we have all helped to sustain.

So much for the scope and complexity of the problem. Let us examine now *one* hypothesis about effecting change in teacher education that has guided the work of a significant number of people, historically those associated with the audio-visual or instructional media area and more recently, content specialists themselves. Briefly, this hypothesis is that the creation of new materials is an effective way of changing the training of teachers. A recent and highly influential argument for such "material-based training" was posed by B. Othanel Smith in his book, *Teachers for the Real World* (1969). In this volume, he defined the need for two kinds of materials: "protocol materials" designed to contribute to the development of competence in interpretation and "training materials" designed to contribute to the development of skilled performance. Even though certain empirical evidence has blurred the distinction between these two kinds of materials as *separate types* (Kleucker, 1974), the essential argument for the necessity of new materials for training remains. In this paper, the author will pose some general propositions about the design and development of materials for training, basing those propositions on both the psychological and practical dimensions of the problem. As the implications of these propositions are explored, a rationale will also be developed for basing innovation in training upon innovation in materials. In fact, considerations of design and rationale will be interwoven throughout.

To a significant extent, the discussion in this paper will draw upon the work of the National Center for the Development of Training Materials in Teacher Education (based at Indiana University) and of several projects associated with the Protocol Materials Development program. Under United States Office of Education funding during the past four or five years, those associated with these projects have struggled valiantly (and sometimes successfully) with the problems of conceptualizing, designing, producing, and evaluating protocol and training materials for a variety of concepts and skills at different grade levels and in different subject areas. As a participant in and observer of these efforts, the author has developed some hypotheses, reservations, hunches, and convictions that might well be sifted for the wheat and the chaff.

Some Matters of Definition

First, we should attend to a few definitions. The term "materials" is used in this paper to refer to the full range of media including, for example, print, motion picture, audiotape and photographic stills. The term "materials for training," on the other hand, is restrictive. It is meant to refer to those materials that are not primarily informational in purpose. Rather, such materials are intended to develop certain functional skills ranging from observational or interpretive to performing. In other words, "materials for training" may be designed to contribute to interpretive competence or to skilled performance or to both. In terms of content, such materials are likely to confront the trainee with the recorded behavior of teachers and pupils. For the purpose of developing interpretive competence, the behavior portrayed should illustrate or exemplify clearly defined concepts about teaching, about learning, or about the substance of a content area.[1] For the purpose of contributing to skilled performance, the teacher behaviors portrayed might usefully illustrate specifiable teaching skills of clearly apparent utility in the classroom. A major practical outcome of this emphasis on the portrayal of behavior is that many materials for training will take the form of such audio-visual media as film or audiotape.

To further clarify the distinction between the informational and training purposes of materials, we might consider an illustration from the field of educational psychology. In the social-psychological area of that discipline, techniques for describing group relationships have long been of interest. One such technique, that of "sociometrics," is a questionnaire-based method for gathering and plotting data on the status of individuals in a group. This easily gathered data yields information by means of which a teacher may locate students in such categories as "social isolate" or "neglectee." Many textbooks in educational psychology treat the sociometric technique in an informational sense. They may even show a completed plotting of data, attach labels to given categories, and discuss the significance of the results. No textbook of which the author is aware, however, actually puts the learner through the process of organizing, plotting, and interpreting "raw sociometric data" although this can be fairly easily done. In terms of the definition in this paper, textual material incorporating the latter characteristics could be classified as "materials for training." It is probably safe to say that most current instructional materials in teacher education are in fact informational in character (including more than a few that are inappropriately classified as "training materials"). It is this emphasis on "information giving" that might lead one to conclude that even many educational films, for example, are not classifiable as "materials for training."

---

[1] Materials that successfully meet this criterion of "concept instancing" are generally classifiable as "protocol materials."

Designing Materials for Training

Let us turn next to the development of materials for training with particular reference to moves and strategies in teaching mathematics. Here a note of caution is in order. The development of such materials presents a number of conceptual, design, and technical problems. The present paper, however, will concentrate on the problems of designing and producing materials for training. The conceptual problems, such as defining and analyzing the concepts or skills to be portrayed, are the business of the content specialist. In terms of this monograph, these are problems for the mathematics educator.

The problem of design is one with which both the content specialist and the technical or media specialist must contend because both the nature of the conceptual area and the nature of the medium influence the design possibilities. Work on this task might well be guided by some very general suggestions from the psychological literature on the acquisition of concepts and skills. In the following section, a few generalizations from that literature will be summarized. In a later section, we will consider some practical guidelines suggested by probable conditions of use in teacher education programs.

Suggestions from Psychology

Three years ago, the National Center commissioned an educational psychologist, Bryce B. Hudgins, to survey the major literature on the acquisition of complex skills and to draw from that literature implications for the design of materials to be used in the development of skilled performance. In this section, we will consider some major points from Hudgins' study; a more thorough analysis is provided in his monograph (Hudgins, 1974). The general recommendations that he makes about the development of complex skills (such as learning to use teaching moves and strategies in mathematics) may not seem very surprising. This is, in fact, a case where the implications of research are supported by a certain correspondence with common experience.

Hudgins identifies three major stages in the development of skilled performance: an overview or informational stage in which the trainee develops a general perspective of the context in which the behavior is to be performed; a practice or performance stage where he or she has the opportunity to exhibit behaviors, within either simulated or actual conditions; and a "feedback" stage in which the trainee gains information about the efficiency and skill of his or her performance. In the event that these stages seem overly general, remember that Hudgins' argument for their effectiveness rests upon two additional and critical conditions: (a) the refinement and precision with which a complex skill is analyzed for the component behaviors to be practiced and (b) the arrangement of extended and intensive rather than occasional or incidental practice. Within this set of generalizations, there may well be implications for the development of skill in the teaching moves and strategies of mathematics. At this point, it is sufficient to note that the teacher trainee, according to this training model, would proceed through (a) an overview of teaching moves within teaching strategies and perhaps an overview of teaching strategies in the context of a lesson, (b) practice both in making moves and using strategies, and (c) some feedback on the skill of his or her performance in both.

Within this general training model, there are also some evident, and less evident but interesting, occasions for the use of materials. In the first stage of the model, for example, the use of protocol films illustrating specified teaching moves or strategies could clearly establish these moves or strategies as definable behaviors and, at the same time, provide an overview for the trainee to gain a general perspective on teaching moves and strategies in the context of classroom teaching. In fact, the rationale for such a protocol film series has been recently developed by Cooney, Kansky, and Retzer (1975). In their very interesting monograph, the authors specify a selection of concepts referring to teacher behaviors, or moves, that are instrumental in teaching specified concepts in mathematics. They have also produced a "rough draft" videotape demonstrating the utility of protocol films in mathematics education.

In the second and third stages of this training model, those of practice and feedback, there is another interesting opportunity for the development of materials. At the National Center, there has been some encouraging practical success with a highly simulated training technique that provides an opportunity for the trainee to respond under modified "feedback" conditions. Working in the general area of developing skill in reacting to pupil responses, several associates of the Center have constructed pilot printed manuals and audiotapes that present bits of classroom dialogue, or interrupted classroom dialogue, with actual teacher reactions either omitted or delayed. The trainee is required to construct, in either written or oral form, an appropriate reaction to each student response. Informational feedback is provided for the trainee through brief editorial comments suggesting one or more aspects of a student's response to consider in constructing a reaction and through providing the teacher's actual reaction after a delay.

Our experience with this technique, along with the successful use of comparable simulated methods by others (see for example Borg, 1975b), suggest that it is a promising training method. It is interesting to speculate about the application of such a simulated technique to the acquisition of moves in teaching mathematics. For example, correct and incorrect pupil responses to specified teaching moves might be filmed, the trainee being required to construct successive or alternative reactive moves in such a simulated form as audio-recording.

Let us return once more to the matter of developing films illustrating specific moves and strategies. As in the case of skill acquisition, there are certain suggestions in the literature on concept acquisition that might provide useful guidelines for design (Clark, 1971; Ellis, 1972; Hudgins, 1972). If a specified concept referring to a teaching move is to be illustrated on a protocol film, for example, the behavior referred to by that concept should initially be isolated as much as possible. This calls for a very brief filmed excerpt with a sharp focus on the relevant teacher behavior. Examples as well as nonexamples of the concept should be contrasted, examples predominating in frequency. The behaviors referred to by the concept should be varied across teachers or pupils and across classroom settings. Ultimately, but not initially, the behavior referred to should be presented in a relatively "noisy" or complex setting. Obviously, this maximizes the similarity between the training condition and the transfer condition (i.e., recognizing behaviors in actual classroom settings). Finally, the relevant attributes or indicators of the concept should be "highlighted" as much as possible.

Once again, of course, these suggestions should be treated as guidelines rather than as prescriptions. In our own work, we have modified them or varied them depending upon considerations pertaining to everything from the dimensions of the specific concept to the aesthetic quality of the film. At one time or another, however, these varied suggestions have all been met in the protocol films on teacher-pupil interaction developed through the National Center and the Indiana University Protocol Materials Project. Viewing these films does provide one or more concrete examples of each of the above specifications (see Gliessman, 1974).

Implicit in our analysis to this point have been several arguments for the development of specially designed materials as a basis for training. These might be briefly summarized. Materials that are produced in an auditory or visual format have the unique capacity of directly portraying behaviors that illustrate concepts or model skills. The expectation is clearly that the use of such materials will contribute to the acquisition of concepts and of teaching behaviors. In the case of concept acquisition, particularly as evidenced by the ability to identify the referent behaviors in recorded teaching episodes, there is a growing body of empirical evidence that confirms this expectation; the systematic use of well-designed protocol materials does lead to concept acquisition (Cooper, 1975; Gliessman, Pugh, & Perry, 1974).

There is also evidence that the use of printed transcripts and of films, in conjunction with more or less highly simulated practice, results in the acquisition of specified behaviors by teachers (Borg, 1975a). Finally, there is evidence that viewing and analyzing specified teaching behaviors on film results in as frequent use of those behaviors by trainees in a simulated teaching setting as does overt practice with feedback (Kleucker, 1974). If such results continue to be replicated for these and other teaching behaviors, the combined effectiveness and efficiency of well-designed materials for training should attract increased attention in teacher education programs.

Suggestions from Practice

When we turn to the problem of instituting or implementing new training procedures in teacher education programs, an equally strong case can be made for the development of materials. In attempting to implement or institute a new procedure, it is instructive to ask exactly what the agency or vehicle of that implementation is to be. Obviously, instituting a training procedure depends ultimately upon knowledgeable people, either specialists in that procedure who assume responsibility for instruction or those already responsible for instruction who become specialists in that procedure. Beyond this, however, the possibilities are clearly limited. The specialist (in this case, the mathematics educator) can, for example, describe a teaching move, show or display that move, or put someone through that move. Ideally, perhaps, he or she would work through some combination of these. Predicting that initially there will be few specialists and many novices with respect to moves and strategies in mathematics, the practical limitations of depending upon such personal intervention as a means of "inservice training" become obvious. At the same time, the usefulness of well-designed material for training becomes clear. The author's hypothesis is that distributable materials that portray teaching moves and teaching strategies can be instrumental in accomplishing the broadest possible dissemination.

It is also instructive to view materials from the perspective of research on teacher training. In his paper, Retzer discusses the importance of "delivery system research." In so doing, he calls our attention to the need for empirical studies on the outcomes of training in teacher education. From this perspective, materials for training may profitably be viewed as components of training or treatment variables (Gliessman &

Ingersoll, 1974). The difficulty of drawing inferences from, to say nothing of replicating, the results of much research in teacher training is partially due to the paucity of descriptions of the conditions of training. Materials for training, being concrete and distributable, can significantly add to the descriptive detail about the training treatment. Such materials can also provide the research worker with added practical control over the conditions of training.

Just as well-designed materials can affect practice, the problems posed by the conditions of practice have clear implications for the design of materials. In this paper, we have already attested to the difficulty of instituting any new procedure in teacher education. Conventions of administration, habits of instruction, and complexities of instructional service are likely to present many points of resistance. What is surprising is how developers of new training procedures, and of materials for training, seem to confound these existing problems of implementation by a lack of foresight in design. A carefully validated set of materials for training is of little consequence if it is so voluminous, so detailed, or so intricate that it is impractical for normal classroom use. It is probably an exaggeration to say that a complex set of materials that is clearly superior to other materials for achieving an intended outcome will go unused because of that complexity. It is conceivable that a new training procedure might be so demonstrably superior in its effects that it would be widely adopted whatever logistical problems it posed. It is much more likely, however, that we are really talking about the creation of materials and procedures that produce results equal to, or perhaps somewhat better than, an alternative set of materials or procedures. Because of this, efficiency in design becomes important.

During the past four years, the author has had occasion to look through a number of newly developed materials for training. The impressic they convey is not generally positive. They are frequently too long, too "home made" in appearance, and too complexly organized to be of practical use. One must remember also that many, probably a majority, of these same materials are unvalidated in terms of their stated outcomes. The general prospect is not reassuring.

More systematic evidence of this kind was recently reported by three National Center associates who studied the probable utility of a wide range of existing materials for training in inservice settings (Ingersoll, Jackson, & Walden, 1975). Ideally, of course, such materials should be of particular use in inservice training because they ostensibly focus on the development of teaching skills. As part of this study, these investigators surveyed the conditions of inservice training (availability of equipment, space, personnel, time, etc.) in 26 school systems of modest size in the midwest; they then compared these training conditions with the conditions apparently assumed by the hundreds of materials catalogued in two major reports (Teacher training products, 1974; Houston, 1973). Briefly, the authors concluded that the actual conditions of inservice training in the schools surveyed would make the use of many of these materials highly impractical. For example, the materials catalogued frequently required too much training time often assumed the continued availability of space that did not exist, and sometimes required a combination of audiovisual equipment that was probably too complex to handle. In addition, these materials commonly assumed a specialized leadership skill that was not readily available to the school systems. Finally, and somewhat parenthetically, many of these materials were actually informational in nature; they serve largely to inform or motivate the user rather than to provide training in a performance.

Producing Materials for Training

The theoretical and practical considerations to which we have been attending are not, however, the only factors that influence the design of materials. The specific medium with which one chooses to work is also a significant factor since the latitude for design is not the same across all media. For example, certain visual effects sometimes used in film (such as "freeze frames" that stop motion and thus, in effect, introduce "photographic stills" into a motion picture) are essentially redundant to filmstrip and irrelevant to audiotape. In any event, the actual production of materials, the point at which a design idea is committed to physical form, is certain to add some new dimensions to that design idea.

It is only fair to point out that the area of production has its share of "traps" for the unwary developer; he or she really does need the guidance of a media production specialist to traverse the area safely. As content specialists, we are very unlikely to have the technical knowledge to handle production problems with finesse and efficiency. Our purpose in this section is simply to raise a few general questions about production for initial guidance and not to completely "map" a complex area.

What media are most effective? Although frequently asked, this question is really unanswerable in a general sense. Even when care is taken to define what is meant by "effectiveness" (in terms of learning outcomes or user reactions, for example), one will not find any decisive data in the literature although empirical studies addressing the question are plentiful (Levie, in press). Among the reasons for this uncertain state of affairs is the fact that any so-called "single medium," such as motion picture film, is highly variable across specific films in design components that may affect learning or viewer response. Thus, comparing one medium with another generally becomes an exercise in comparing undifferentiated or poorly described media forms.

In short, one is not likely to find convincing general evidence on effectiveness that will lead to the selection of one medium over another. A more promising approach has been described and rationalized by Levie (in press). Essentially, his is an analytic approach that calls upon the developer to consider the interrelationships among the learner, learning task, and learning environment in terms of certain critical characteristics that are shared by the medium. At a minimum (and here his dimensions will be simplified for the sake of clear communication), the developer might analyze the skill to be learned for critical characteristics that are also reflected in specific media forms. For example, if he or she decides that a specified teaching move has certain critical gestural characteristics, videotape or motion picture might wisely be selected as the medium because each has the capacity to communicate visual motion.

Although there is apparently some evidence emerging on the effects of different media characteristics, it is probably wisest at this point to view the question of the effectiveness of a specific set of materials as one for the developer of those materials to assess through carefully designed evaluation studies (see, for example, Gliessman, Pugh, & Perry 1974). Once the spectre of "comparing media" is exorcised, the developer may even decide to vary the design or the components of that set of materials to evaluate the effect of such variation on learning outcomes. In this way, he or she may begin to generate more generalizable knowledge about the design of effective materials.

What are the comparative costs of media production? In spite of the logic of this analytic approach to media selection, the author's impression is that most developers have a bias toward working with certain media rather than others. Of course, one can have more or less expensive biases, and that is the theme of this section. It is clear that the cost range of producing in different media is very great. Working through the same producer (who is exercising consistently high standards) and recording the same classroom episode, one can spend from as little as $400 for a finished, ten-minute audiotape to as much as $12,000 for a finished, ten-minute motion picture in color with sound. Generally speaking, then, with technical standards and content held constant, it is more costly to produce on videotape or motion picture film than on audiotape or 35mm transparencies.

What is equally important to understand, however, is that the cost range within any medium can be as great as the cost range between media. For example, depending partly upon whether a teacher and student in a tutorial session or a teacher and a total class is being filmed, one can spend from as little as $200 per minute to as much as $1200 per minute of finished film. Thus, in a real sense, most generalizations about comparative costs of producing in different media are probably specious; the conditions and standards of production significantly influence the cost of producing in any medium.

Is it really critical to meet high technical standards in production? The answer to this question depends upon the phase of development. If a developer is producing first approximations or "rough drafts" of an anticipated product, high technical standards should not be a primary concern. In fact, much of the development work at the National Center has routinely involved the production of such rough drafts during the formative stage of development and evaluation. For example, prior to motion picture production, film outlines or scripts are initially committed to "home made" videotape production. At this point, high quality sound and picture is not a concern; content and message quality, on the other hand, is. In fact, attending to high technical standards at this early point in development is probably counter-productive since the emphasis should be on making needed changes in content and design. Re-shooting or re-editing finished film is very expensive; revising "home recorded" videotape is not at all expensive.

In the final stages of development, and particularly when a decision has been made to distribute materials to other users, high technical standards must be met. It is an illusion to argue that conceptual elegance or high content validity will make up for such technical inadequacies as an inaudible sound track or an obscure picture. Materials distributed beyond the local level are almost inevitably in competition with a large volume of "polished" curricular and instructional materials. What is less obvious is that, in distributing materials, one is also in competition with a lifetime of media viewing experience by students who have lived with commercial television and commercial motion picture. Although the conceptual standards of much of commercial television or film might legitimately be questioned, there is much less question about its technical adequacy.

What is the best way to generate and record examples of teaching moves? There is an assumption in this question that should first be made explicit. This assumption is that audio or audio-visual examples of teaching moves are central to the development of materials for training in mathematics education. In other words, this question assumes the creation of protocol materials illustrating teaching moves and, perhaps, teaching strategies. Previously in this paper, we have noted the plausibility of such a format in the creation of useful materials.

The question can be addressed most productively by rephrasing it to ask what are the most efficient procedures for generating and recording examples that (a) are unambiguous referents of clearly defined concepts (let us call this characteristic "referential validity") and (b) are behaviorally authentic, rather than contrived or artificial (let us call this "behavioral validity"). With these characteristics in mind, we might assess several approaches to producing tape or film footage.

One very common approach is to prepare a complete script of an episode that is then enacted and recorded on tape or film. For example, one might write and perform a script of a classroom episode incorporating instances of specified teaching moves. By this approach to production, one can achieve considerable efficiency and referential validity (since the examples of teaching moves are preplanned and specifically written into the script to be "acted out"). However, this approach leads to problems of behavioral validity. In the first place, in the preparation of a script, it seems uncommonly difficult to construct an array of illustrative behaviors having the range and variation of behaviors occurring naturally during the course of actual instruction. Even if this problem is solved (as it might be, for example, by drawing examples from extended classroom observation), the best prepared script is still ultimately only as effective as the performers who enact it. Unless these performers are highly skilled (and, in the author's own experience, this usually implies that they have professional skills as actors or actresses), a certain degree of artificiality or "stagedness" generally characterizes the performance.

A contrasting approach is to "document" or record unscripted and unstructured behavior occurring as a "natural" part of a conventional instructional episode. More accurately, the only structure provided would be that content and organizational structure exercised by the mathematics teacher as part of his or her normal classroom instruction. Teaching moves would thus be "captured" on film as they occurred during the course of instruction. This approach should increase the degree of behavioral validity once the teacher and students become accustomed to the presence of the recording equipment and production crew. However, the referential validity of the examples that do occur is likely to be a problem; "naturally occurring" teacher behaviors are frequently so vague, complex, and confounded that they do not illustrate the pertinent concepts with sufficient clarity and salience for training purposes. Finally, this type of unstructured approach is notoriously inefficient in that it requires the recording of an excessive amount of footage to obtain a few "clean" examples.

A third approach was used in producing the film footage for the protocol film series, Concepts and Patterns in Teacher-Pupil Interaction (Gliessman, 1974). Since that series focuses on concepts referring to teacher behavior, some direct applications of this approach can be made to the problem of exemplifying teaching moves. Using what is best called a "structured documentary" approach to filming, three elements or components of each episode were prearranged.[2] First, the "film teacher" was carefully trained in specific teaching behaviors that should occur with considerable emphasis during the episode. Second, the specific content of instruction was predetermined in consultation with the film teacher; this allowed for the selection and development of content that was most appropriate for the critical teaching behaviors. Third, the general mode of teaching (in this case, teacher-led class discussion) was established, again to provide the context for the critical teacher behaviors to occur with sufficient salience and frequency. Students in the film were instructed prior to production only in the general academic content of an episode. During filming, they were free to respond to the film teacher spontaneously. Having no script, the film teacher was similarly free to develop and conduct the classroom dialogue (in other words, to "teach naturally"). During filming, the teacher's performance was "monitored" for occurrence of the critical behaviors; reshooting was done when necessary.

---

[2] Here credit is due Michael Waddell, Promotional Marketing Service, Louisville. As film producer for this series, he was instrumental in conceptualizing and successfully executing this filming strategy.

From this general footage, behavior excerpts were later drawn that most clearly and cleanly illustrated the specified concepts. Drawing in this way upon filmed episodes that had been selectively structured and carefully monitored resulted in a high degree of referential validity. Such initial structuring also increased the efficiency of filming because unusually large amounts of footage did not have to be recorded to obtain examples of the critical teacher behaviors. Finally, the latitude allowed by the absence of a script resulted in the kind of spontaneity that contributes to a high degree of behavioral validity. Thus, such an approach to filming has specific advantages that recommend it as an effective means of generating and recording moves in teaching mathematics.

Summary

The thesis of this paper is that change in the training of mathematics teachers can be effected to a significant extent through the creation of new materials for training. More specifically, the skillful performance of specified teaching moves and strategies can be developed through training procedures based on specially designed audio and visual materials. Since such training involves both conceptual and performance elements, implications for the design of materials were drawn from the empirical literature on both concept and skill acquisition. Since such materials must be of practical value in training, implications for design were also drawn from educational practice. Finally, questions about the actual production of materials were explored: the comparative effectiveness and cost of different media, the importance of technical quality, and the most effective production strategy.

References

Borg, W. Protocol materials as related to teacher performance and pupil achievement. Journal of Educational Research, 1975, 69, 23-30. (a)

Borg, W. Moving towards a breakthrough in teacher education. Education, 1975, 95, 302-323. (b)

Clark, D. C. Teaching concepts in the classroom: A set of teaching prescriptions derived from experimental research. Journal of Educational Psychology. 1971, 62, 253-278.

Cooney, T., Kansky, R., & Retzer, K. Protocol materials in mathematics education: Selection of concepts. Acquiring Teaching Competencies: Reports and Studies (No. 7). Bloomington, Indiana: National Center for the Development of Training Materials in Teacher Education, 1975.

Cooper, J. E. A survey of protocol materials evaluation. Journal of Teacher Education, 1975, 26, 69-77.

Ellis, H. C. Fundamentals of learning and cognition. Dubuque, Iowa: William C. Brown, 1972.

Gliessman, D. Concepts and patterns in teacher-pupil interaction: A protocol film series. Bloomington, Indiana: National Center for the Development of Training Materials in Teacher Education, 1974.

Gliessman, D., & Ingersoll, G. M. Educational psychology and the instructional media. Viewpoints, 1974, 50, 21-36.

Gliessman, D., Pugh R. C., & Perry, F. Effects of a protocol film series in terms of learning outcomes and reactions of users. Bloomington, Indiana: National Center for the Development of Training Materials in Teacher Education, 1974.

Houston, R. W. Resources for performance-based education. Albany, N. Y.: Multi-State Consortium on Performance-Based Teacher Education, 1973.

Hudgins, B. B. The portrayal of concepts: An issue in the development of protocol materials. Acquiring Teaching Competencies: Reports and Studies (No. 1). Bloomington, Indiana: National Center for the Development of Training Materials in Teacher Education, 1972.

Hudgins, B. B. Self-contained training material for teacher education: A derivation from research on the learning of complex skills. Acquiring Teaching Competencies: Reports and Studies (No. 5). Bloomington, Indiana: National Center for the Development of Training Materials in Teacher Education, 1974.

Ingersoll, G. M., Jackson, J. H., & Walden, J. Teacher training needs, conditions and materials: A preliminary survey of inservice education. Acquiring Teaching Competencies: Reports and Studies (No. 8). Bloomington, Indiana: National Center for the Development of Training Materials in Teacher Education, 1975.

Kleucker, J. Effects of protocol and training materials. Acquiring Teaching Competencies: Reports and Studies (No. 6). Bloomington, Indiana: National Center for the Development of Training Materials in Teacher Education, 1974.

Levie, W. H. How to understand instructional media. Viewpoints, in press.

Smith, B. O. Teachers for the real world. Washington, D. C.: American Association of Colleges for Teacher Education, 1969.

Teacher training products: The state of the field (Research and development memorandum No. 116). Stanford, California: Stanford Center for Research and Development in Teaching, 1974.

Reflections from Research:
Focusing on Teaching Strategies
from Various Directions

Marilyn Suydam

The Ohio State University

Introductory Reflections

The search for and the control of teaching strategies has, at one and the same time, idealistic and pragmatic goals. The promise of and progress in the development of a theory of instruction is coupled with the chance of providing some pertinent evidence on behaviors which will help teachers teach more effectively. Too often, research has focused on the pragmatic need for an answer to a specific question in a finite amount of time. That the idealistic need should also be served by each separate piece of research has been ignored. The pressure from teachers to "give me something I can use tommorrow" has concentrated attention today on materials and activities. How to decide when to use them--and how to teach--at times seems to have become irrelevant to many. Yet, in the past decade an increasing number of teachers have become interested in research, and in what research can tell them about effective ways of teaching.

Compare the reflections of teachers at two levels; the similarity of their need is obvious.

> I am a teacher.
> I teach, and children learn, and I hardly know why.
> I try to figure out what I do as I teach that makes
> children learn, but I can only see some vague clues--
> but there is no pattern.
>
> I am a teacher educator, and I suddenly realize that I know
> virtually nothing about the act of teaching.
> Oh, I can teach.
> I can show others how to teach.
> I can even tell a few things to do that probably will
> make others more effective teachers.
> I can recognize good teaching, and know some of the things
> that made that teaching seem good--and I can recognize

> poor teaching, and know some of the things that made it poor. But I feel as if I'm playing a game for which I don't know the rules.

As we rationally explore teaching strategies and aim for a theory of instruction, let us not forget that teaching proficiency is acquired in many ways. The ability to teach, the capacity for teaching, seems born in some: They do the "right" things naturally, constantly varying their response to the stimulus that is the learner. Others learn by watching, by modeling behaviors, by modifying what they have seen others do. And others struggle just to use the "bag of tricks" they have been handed in a teacher education program. All can profit from seeing the patterns of teaching, patterns that are collections of teaching strategies.

Throughout the years, attempts have been made to identify the characteristics of an effective teacher. There is some consensus that it is important for a teacher: (a) to like children, (b) to communicate with learners, (c) to know what to teach and to like that content, (d) to have a philosophy of teaching, and (e) to understand how to teach.

But for years, researchers in mathematics education (as well as other curricular areas) have seemingly been avoiding the main point as they have attempted to ascertain what makes an effective teacher. They explored the teacher's background, the number of mathematics courses the teacher had taken, the age of the teacher, the number of years of experience the teacher had, and a host of other factors, to determine correlates of teaching effectiveness. Certain correlates were indeed found, such as warmth and enthusiasm and managerial ability. But such correlates have little promise as variables for controlled experimentation on the relationship between teacher behavior and student achievement. Despite the fact that in study after study limited correlations or few useful ones were found, the pursuit went on.

Thank goodness we are no longer so myopic. We may never find strong correlates of achievement: Perhaps the process of teaching is too complicated for that. As Fey (1969b) noted:

> The question of predicting teacher effectiveness is not simply answered by direct measurement of obvious variables, but must be viewed as a complex interaction of several interrelated classes of variables. (p. 541)

At long last we are looking at the process: What is it that does go on in a classroom, in a lesson, in a moment of teaching? What is it that the teacher does which contributes significantly to the student's learning? Analyzing what actually goes on in the classroom, to find clues about what effective teaching is, seems so utterly realistic that it is amazing that so much time was wasted in not doing just that. Perhaps we will find out more about how to teach teachers to teach, with the ability to understand why certain teaching procedures work and the capability of generating more effective strategies. Perhaps we will find out how to predict what will happen under specified conditions, so that we can select strategies that will ensure learning. (However, we must also keep in mind as we observe what goes on in the classroom that what teachers do may not be what teachers should do. For instance, Brown (1974) reported a very heavy dependence on the textbook by the 42 algebra teachers and their classes whom he studied. Mathematics became a sterile sequence of homework, discussion, and new homework.)

While the search for what makes a teacher effective pursued its course into the teacher's background, other research with a different goal, ignoring the specifics of the teacher-learner exchange, explored comparatively more general areas: what content to teach, whether to use this algorithm or another, or the effects of alternative methods. Teaching strategies were implicit in each, but usually only generally or globally defined. Many times, in one way or another, it was "concluded" that the teacher made the difference, that almost any alternative procedure is effective when employed by a good or effective teacher. (What a good or effective teacher is, however, was undefined.)

We have confirmed many things with these studies. We can point to research and indicate to teachers certain things that research has made evident. Some are specific; for instance, algorithm A promotes achievement better than algorithm B. We have also confirmed the effectiveness or appropriateness of many procedures which can be generalized almost as "rules":

1. Plan systematically.
2. Base instruction on the readiness of pupils.
3. Group pupils and have pupils teach each other.
4. Teach with meaning.
5. Teach for transfer.
6. Motivate and praise.
7. Provide practice following understanding.
8. Spend at least half of the class time on developmental activities.
9. Use concrete materials before proceeding to abstraction.
10. Provide periodic review and spaced practice.
11. Diagnose errors and provide instruction on the basis of diagnosis.
12. Teach a variety of problem-solving procedures.

These obviously do not combine to form a theory of instruction. In fact, only recently has there been serious recognition of the need for a theory of instruction and a serious attempt to develop and find evidence to support a theory of instruction. Among the various attempts in mathematics education to develop such a theory, are two which are attacking the problem from different directions. The work of Heimer and his students (Heimer & Lottes, 1973) is an effort to define sequences of instruction and the variables related to sequences which affect learning, generally at the elementary-school level and highly related to learning outcomes such as transfer and retention. Our concern is with the other instance, the work of Henderson and his students who define strategies of teaching and variables related to teaching strategies which affect learning, generally at a post-elementary-school level. Mention is made of the existence of the two approaches in order to point out two facets which they have in common. In both instances, a model for the research has been rationally developed, and the attempt to confirm and expand portions of that model is being made through a set of coordinated studies. These two facets alone are firm steps forward in educational research.

Dossey (in this monograph) provides a concise presentation tracing the development of relevant work on teaching strategies as defined by Henderson's model for teaching concepts. He gives an excellent summary of the status of thought and research on the topic. He has excluded some studies not relevant to his purposes (e.g., Cooney, 1970; Wolfe, 1969) in order to focus on those studies which have tested the utility of Henderson's models. Dossey's questions are valid ones; reflections of them will arise elsewhere in this paper.

The extension of the model he proposes holds some promise for coordinating future research efforts. The model needs further specification, for example, specific content and/or grade/age levels need to be incorporated. If research is to be effective at developing a theory, each study must be defined precisely in terms of its relationship to any other study based on the model. Merrill and Wood (1974) carry this one step further; they proposed that every study should be related to every other study through clear specification of each component. Coordinated planning must incorporate all aspects of a model, and a model must explicitly incorporate all aspects of concern in research.

In addition to further research, there is also a need for (exact) replication of studies. Dossey makes this evident as he attempts to compare the Rector and Henderson (1970) and Dossey (1972) findings. The strategy not used by Rector accounted for the differences in Dossey's study. There is no way of determining the meaning of this difference or the stability of any findings unless a study is conducted which parallels a given study. He also attempts to compare the findings of Rollins (1966) and those of Gaston and Kolb (1973). Too many variables are being changed at once. In the attempt to "prove" the model or theory as quickly as possible, we must not lose sight of the need to interpret and interrelate findings. There are two distinct aspects of research related to strategies, and we need information on each: (a) What goes on in the classroom between teacher and student as specific strategies are manipulated and (b) is there a relationship between what can be observed and pupil achievement?

There are literally a horde of variables which must eventually be considered. These variables must be built into the model in some way and explored systematically. Many are noted in other papers of this monograph:

for instance, timing or pacing of instruction, ability level of students, characteristics of students (e.g., learning style), and type of content. There are many questions to consider as the research model is expanded and as each individual study is designed. Not all questions can be answered, but the problems implicit in their formulation must be recognized.

Reflections from Research

It is frequently true that a research study will raise more questions than it answers. Let us consider some studies which have been selected because they provoke some questions which should be considered in connection with research on teaching strategies. All are drawn from the set of studies on mathematics education. Quality was not the determining factor: Some studies have obvious "faults," and may have been selected for precisely that reason. Each study and/or the findings from it are briefly described to define the setting from which the question arises. Hopefully, the source of and/or the reasons for the questions will thus be apparent. Some of these questions are stated for emphasis of points noted elsewhere in this monograph. Some point out the need or advisability of considering certain variables which are not explicated in other papers. Some cannot be answered by research, while some could lead to researchable problems. There is no attempt to raise all of the questions which might arise from a cited study; rather, one or two which might cause the reader to develop others are posed.

Not unexpectedly, many of the studies cited involve some form of interaction analysis (e.g., studies 1-12). As researchers have tried to analyze what goes on in the classroom, various instruments or procedures have been developed and used. Until systematic procedures were devised, in fact, attempts to analyze what happened in the classroom were rather confusing. Research with these instruments has added to our knowledge of classroom occurences, though it must be kept in mind that what is "found" is a reflection of what an instrument or procedure allows to be recorded.

Since questions are one of the valued techniques for teaching, evidence on what types of questions teachers ask is useful. Some studies of teacher-pupil interaction which focused directly on the type of questions are cited (e.g., studies 13-17).

One study (18) uses a computer to simulate a teaching incident. Studies related to Henderson's model have moved from use of programmed instruction as the teaching mode to study of the teacher in the classroom. Use of either creates certain problems. With programmed instruction, there is precise control of the strategies, but the interaction is also controlled -- and limited. With the actual teacher in a classroom, that very interaction can mean less control of the strategy being used. CAI is proposed as an alternative, or as an intermediate stage between programmed instruction and "live," interactive classroom instruction.

Embedded in the teaching strategies proposed by Henderson, and explicated more extensively in some studies (e.g., studies 19-21), is the use of nonexamples and counterexamples. They point out a particular problem, one of definitions. A lack of agreement on definitions or use of terminology plagues much research. The meaning of the terms "nonexample" and "counterexample" is presented by Henderson (1967).

> Nonexample: An object which is not a member of the referent set (that is, not an example) is designated. Usually this move is employed when experience has shown that students make errors based on not knowing certain necessary conditions.
>
> Counterexample: An object is named (or otherwise designated) that falsifies a generalization purporting to characterize the members of the referent set. This move is used after a characterizing move and is never the first move in a concept venture. It is often used to enable a student to correct a misconception. (p. 576)

These definitions may be clear, but all researchers have not made this distinction in their use of the terms. To confound the situation more, a third term "negative instances" is also used, probably as a synonyn for nonexample, but possibly as a generalized term to cover both nonexamples and counterexamples. This makes research on this facet somewhat difficult to interpret, particularly in cases where specific illustrations are not included in a research report.

Now let us consider the selected studies and some questions.

Study 1: Options from Organization

With a statement about the nature of teacher-student interaction, Hudgins and Loftis (1966) described the options of students and teachers in the classroom:

> Interaction in the classroom occurs under special conditions that set it apart from interaction in other kinds of settings. In most cases the teacher originates the interaction. The teacher has the prerogative of raising questions, setting tasks, and evaluating behavior independently of the desires of pupils. In the arithmetic class the teacher is free to demand interaction with any pupil at any time, but the converse is not altogether true. Usually pupils interact with the teacher only when, in effect, the teacher signals that such interaction is permissable. While interaction in the classroom can be initiated either by the teacher or the pupil, it is important to remember that pupil-initiated interactions are contingent upon teacher sanction. (p. 146)

In twelve fifth- and sixth-grade classrooms, Hudgins and Loftis studied 11 "visible," 11 average, and 9 "invisible" students. ("Visible" and "invisible" students were high in arithmetic achievement, but differed in the extent to which peers recognized that ability.) It was found that the students did not differ in the average number of interactions with the teacher or in the feedback they received from the teacher. The teachers appeared to be remarkably consistent in distributing interactions equally among students and in evaluating student responses. Visible and invisible students initiated interaction more frequently than the average students did, but teachers tended to compensate for this difference by initiating interaction more frequently with average students than with the visible or invisible students.

Questions: How does the way in which a classroom is organized affect the type and quality of teacher-student interaction? What type of organization is implicit in each Henderson-oriented study on strategies?

Study 2: Teacher Awareness

Strickmeier (1971) attempted to (a) describe and compare patterns of teachers' verbal behaviors in seventh-grade mathematics classes grouped by ability and (b) determine if seventh-grade mathematics teachers have different perceptions of their verbal behaviors and different expectations of student behaviors for classes of different ability levels. He had teachers complete a questionnaire concerning their expectations of behavior within each class, and then he observed each teacher three times with a high-ability class and three times with a low-ability class (using the OScAR-5V).

Analysis of the data revealed that the teachers did have different perceptions and expectations for the classes of different ability levels. But despite these differing perceptions and expectations, teachers' behaviors were not different for the classes at different ability levels.

Questions: What relationships do and/or should exist between teachers' expectations and their choices of strategies? What student characteristics should be considered when selecting samples?

Study 3: More on Teacher Expectations

Kester (1969) found by observation that seventh-grade teachers communicated with their allegedly bright pupils in a more friendly, encouraging, accepting manner. As the pupils' positive communication to the teacher increased, the teachers' communication to the pupils tended to be positive. Teachers also spent more time communicating with them. Nevertheless, pupil achievement was not significantly affected.

Question: *How does the way in which a teacher views each student affect the strategies for and responses of that student and the class?*

Study 4: Training for Interaction

Smith (1971) worked with a group of preservice elementary teachers. They defined evaluation and discussed its purposes, formulated guidelines for observing classes, observed and discussed filmed lessons, cooperated in the development of an observation instrument, and independently used it while observing six additional films of elementary mathematics teaching. Then they were observed while teaching, using the Adapted OScAR-(EM). Smith reported a greater frequency of (a) use of identified vocabulary, (b) pupil-teacher interaction, and (c) total strategies. Smith concluded that these could indicate that the treatment experiences might have carried over into the classroom.

Questions: *What types of experiences should be given to teachers to make them aware of (a) the role of teacher-pupil interaction and (b) how to develop effective teacher-pupil interaction? Might it be that what is important is not the particular technique used with teachers, but the development of an understanding of teacher-pupil interaction and of teaching itself?*

Study 5: Effect of "Effectiveness"

Dimeolo (1969) reported that students in grades 3 through 6, with highly effective teachers, did not differ significantly in their mathematical achievement from students with less effective teachers. (Teacher effectiveness was determined by differences in teaching patterns, as measured by the OScAR, which were related to pupil gains in mathematical computation.)

Question: *What variables and/or definitions affect a research study so that it appears that students achieve equally well with highly effective and less effective teachers?*

Study 6: Student Participation

Robitaille (1969) provided evidence that the effective mathematics teacher at the secondary level seeks to increase the level of student participation in the lesson significantly more often than does the less effective mathematics teacher.

Question: What is the relationship between amount of student participation and quality of student participation?

Study 7: Time Allocations

Stilwell (1968) studied the behavior of twelve inservice secondary teachers during problem-solving activities in a geometry classroom. A sixteen-category instrument was developed, with ten categories for teacher-talk, four for student-talk, one for structured silence, and one for non-classifiable activity. Analysis of the data showed the following: (a) teacher-talk consumed approximately three times as much time as student-talk; (b) less than three percent of all time in problem-solving involved the activity described as "method for solving a problem"; (c) approximately eight percent of all time was coded as structured silence; (d) looking back at the solution or looking ahead to its implications consumed approximately seven percent of the time, and (e) behaviors in only three of the twelve classes observed differed significantly when engaged in review or in introduction of new content.

Question: What percentage of time is "appropriate" for mathematics teachers to spend in each category for a given type of mathematics activity?

Study 8: Classroom Climate

Vayda (1968) had fourth-graders work on problem-solving tasks. To analyze group performance, he used Withall's Social-Emotional Climate Index, which categorizes teachers' verbal statements and classifies them as predominantly learner-oriented or teacher-oriented. The classes of teachers who were learner-oriented (a) demonstrated more effective group planning, (b) demonstrated greater autonomy in conducting planning discussions, (c) formulated more precise plans of action, (d) were more efficient and successful in solving the problem, and (e) had greater congruence between plans and actual approach used.

Questions: How does the social-emotional climate affect the teacher's choice of strategies and the students' responses? Should climate be considered in the model?

Study 9: Moves and Strategies

Fey (1969a) developed an instrument to describe both pedagogically and mathematically significant components of teacher-student interaction. Four sessions each of five different classes were tape-recorded and the recordings were transcribed. Each transcript was partitioned into a sequence of moves which were then described according to source; pedagogical purpose (structuring, soliciting, responding, or reacting); duration; mathematical content; mathematical activity (developing, examining or applying a mathematical system); and logical process (analytic, factual, evaluative, or justifying). He found that: (a) each teacher spoke more than all of his students; (b) each teacher dominated the pedagogical functions of structuring (80 percent of these moves), soliciting (95 percent), and reacting (85 percent), leaving responding as the major student activity; (c) over 50 percent of all moves were statements or questions of facts, 25 percent were evaluations, and the remaining moves were divided between justifying and analytic process; (d) content emphasis in all classes followed closely the sequence of the textbook chapter being studied; and (e) teacher influence in shaping the direction of classroom activity differed from class to class, but the difference was primarily one of degree rather than kind.

Question: What is the correspondence between Fey and Henderson-oriented studies on (a) definitions of moves and strategies and (b) numbers of moves and strategies observed?

Study 10: Grade/age Level

Mahan (1971) used a system of verbal analysis based on Smith and Meux's (1970) categories, adapted to suit the level of kindergarten discourse. Eleven basic moves were included: Characteristic, Classification, Analysis, Analogy, Differentiation, Instance Comparison, Instance Production, Positive Instance, Substantiation, Negative Instance, and Non-codable. Analysis of tape recordings revealed 1194 interactions in the 16 lessons taught by four student teachers. Descriptive language accounted for 45 percent of the interactions. Comparative actions were least used (11 percent). Instantial moves were used in 41 percent of the discourse.

Patterns were based on blocks of nine sequential moves and identified according to seven types. The mean was 33 patterns used by each student teacher. Type V, a combination of Descriptive and Instantial actions, was most frequently used.

Questions: How do strategies used by teachers at different grade levels differ? What is the role of patterns in relation to strategies?

Study 11: Balance of Moves

Bellack's (Bellack, Kliebard, Hyman, & Smith, 1966) scheme was used by Gordis (1973) to describe the interaction in four first-grade classrooms during instruction on serial ordering. The teacher dominated the discourse by making two-thirds of all moves as well as most of the initiatory moves, 89 percent of the soliciting moves, and 95 percent of the structuring moves. Pupils made 92 percent of the responding moves and teachers 74 percent of the reacting moves. It was also noted that simpler cognitive actions were more likely than more complex ones to be formulated in operational language.

This latter finding was interpreted as giving evidence that evaluation of language usage could reveal the developmental level of children's thought.

Questions: What factors can change the percentage of moves in each category? How do language and content interact with teaching strategy?

Study 12: Language Patterns

Gregory's (1972) major objective was to determine the relationship between the frequency of use of the language of conditional logic by mathematics teachers and their seventh-grade students' conditional reasoning ability. He audiotaped each of twenty teachers' classes five times and administered a reasoning test to students at the beginning and end of the semester. The teachers were ranked on the basis of analysis of the frequency of their conditional moves, that is, how often they used "if-then" language in their teaching. Students of teachers who more frequently used such langauge outperformed students of teachers who made fewer such statements. Thus teachers, through the use of logical language in a variety of situations, apparently helped students to develop greater achievement in the aspect of logic considered.

Questions: Does having teachers focus on specific types of language help students in achieving certain educational goals related to that language? What language patterns do teachers use? What is the effect of these patterns on students' performance? (As one example, if some teachers use many nonexamples, do their students tend to use more nonexamples in their own thinking, or when discussing mathematics?)

Study 13: Level of Questions

Meckes (1972) studied teacher-pupil interaction and teachers' questioning patterns for mathematics in grade 6. A tape recording was made of one class session conducted by each of 100 teachers. Ten-minute segments of each tape were analyzed, and all teachers' questions were transcribed from the 100 tapes and classified into one of seven categories in the Taxonomy of Mathematical Abilities. Meckes stated:

> The results obtained from the Flanders Interaction Analysis appear to indicate that the role of the mathematics teacher has not changed from that of giving information to that of guiding learning experiences. This conclusion was supported by the following evidence: The teacher spent 61.5 percent of the time talking. Direct influence accounted for 50.2 percent of the teacher talk. Although influence amounted to 49.8 percent of the teacher talk, the largest portion of this was in the questioning category. Since

> most of these questions were very narrow, they provided little opportunity for students to express their own ideas. (p. 4246)
>
> Although one of the primary objectives of new mathematics is to foster a spirit of inquiry and to develop creativity, only .5 percent of the total questions were placed in the synthesis category. The two low cognitive level categories accounted for 79.5 percent of the questions asked. (p. 4246)

Question: Could/should the level of teacher questions be incorporated as a dimension of the teaching move or strategy?

Study 14: More on Levels

Friedman (1973) attempted to develop a system that would describe the extent to which teachers seek two important goals of a geometry course: eliciting high levels of student thought and learning the nature of proof. He also investigated the relationship between a teacher's questioning behavior and the performance of his students on a test designed to assess understanding of the nature of proof and the ability to think on various intellectual levels.

Thirty-five lessons taught by 15 teachers were tape-recorded as a geometry theorem was being taught. The coding was by teacher's question and by mathematical activity. The median percent of questions at the memory level was 23, considerably less than the percentage typically asked by teachers of other subjects. Comprehension questions were asked more frequently than other levels of questions; the median was 56 percent. The median percent of application questions was 18. Of 1841 questions asked, only four were higher-level questions, and only twenty involved the nature of proof.

The frequency of the application questions a teacher asked in class was positively associated with student performance on test items at the applications level. For the other types of questions, however, the number of questions a teacher asked appeared to have no clear relationship to student performance on test items of the same cognitive level.

Questions: How does the level of question interact with the specific content being taught? Is there a relationship between the findings of studies like this one and a study on teaching strategies such as Swank's (see Swank's paper in this monograph), in which test items were at different levels?

Study 15: Content Effect

Kysilka (1970) observed 24 teachers (six each in eighth- and eleventh-grade mathematics and social studies) four times, recording their verbal behaviors with the OScAR-5V. Among the findings were:

1. Mathematics teachers asked more convergent and fewer divergent questions, used more procedural-positive questions, and made more directing and describing statements than did social studies teachers.

2. Mathematics teachers talked significantly more than did social studies teachers.

3. The proportion of pupil-initiated statements to teacher statements was significantly greater in social studies classes than in mathematics classes.

Questions: How much control does (or must) the content exert on the patterns of interaction used by a teacher? Are there some behaviors which teachers exhibit in other content areas which could and should be applied to the teaching of mathematics?

Study 16: Training for Questioning

In a study with 30 science and mathematics teachers, Adhikary (1973) gave the experimental group an instructional program requiring the students to work through a programmed instruction unit on classifying teacher questions, discussing the use of different types of questions, writing questions in lesson planning, and discussing reasons why students do not answer teachers' questions. The teachers' question-asking behavior was changed following this instruction. They used more convergent, divergent, and evaluative questions, and fewer cognitive-memory and managerial questions than did those who had not had the instruction.

Question: How does instruction on question-asking interact with instruction on teaching strategies?

Study 17: Effect of Microteaching

Nisbet (1974) used microteaching with audiotaping to help secondary mathematics teachers significantly increase their percentage of use of application and analysis questions.

Question: To what extent can microteaching experiences be used to improve the number, type, and quality of moves and strategies a teacher uses?

Study 18: Training Via CAI

Flake (1974) explored the feasibility of using interactive computer simulations to sensitize preservice secondary mathematics teachers to various questioning strategies. Programs were developed for lesson planning, Henderson's moves and strategies, Polya's approach to problem solving, and a simplified learning theory. Nineteen of 25 students increased their behavior of modeling the prescribed problem-solving strategy; 22 demonstrated an increase in going beyond the first responses of a student.

Questions: What are the advantages of using CAI over (a) programmed instruction and (b) actual classroom observation? Is it plausible to consider use of CAI as an intermediate stage between the other two? Does the use of CAI result in findings that do not apply when a teacher is using the same strategies?

Study 19: Counterexamples for Justification

Wolfe (1969) investigated the verbal activity of justification as it is carried out in the classroom by secondary mathematics teachers. Recordings of 23 class sessions of 11 teachers were analyzed. Eight strategies were identified: six strategies of validation and two strategies of vindication. A finding from the data is that the use of counterexamples to disprove a universal generalization was noted in six percent of the ventures. This may be compared with the use of other types of justification: subsuming generalization, 33 percent (the most frequently used justification in algebra); deductive proof, 23 percent in algebra, 38 percent in geometry (the most frequently used justification in geometry); one or more supporting instances (the most frequently used justification in general mathematics); and pragmatic reasons, 12 percent.

Question: How does this reflection of the actual use of counterexamples reflect the relative importance of their role?

Study 20: Importance of Nonexamples

Sheppard (1972), in a study with 160 fifth-graders, reported that (a) giving divergent examples was superior to giving convergent examples and (b) giving matched nonexamples was better than giving nonmatched nonexamples. Examples were divergent if all three irrelevant attributes were varied and convergent if one irrelevant attribute was varied. Nonexamples were matched if one or two attributes were varied and unmatched if four or five attributes were varied. The dependent variable was the comparison of the student's generalization-discrimination pattern to four predicted patterns.

Questions: How does this specification of type of example compare with the definitions of other research? How content-specific are the findings on the use of nonexamples?

Study 21: "Negative Instances": A Confirmation?

Shumway (1971) designed a study to determine whether an extensive treatment of counterexamples ("nonexamples" using Henderson's (1967) definition) in the development of certain mathematical concepts in grade 8 would result in significant differences in mean scores on tests of (a) general mathematics achievement, (b) specific mathematics achievement, (c) inductive reasoning, (d) syllogistic reasoning, (e) reading mathematical definitions, and (f) tendency to overgeneralize. The study was conducted for 65 class periods with 84 students in four classes; content included quadrilaterals, exponents, and operations. The experimental treatment contained an equal number of positive examples and counterexamples (nonexamples). The control treatment contained only positive examples. The use of counterexamples (nonexamples) was found to have a significant effect on students' tendency to overgeneralize the properties of operations.

Shumway (1972, 1974), using CAI as the instructional mode, has also found that the use of negative instances (nonexamples) as well as positive instances (examples) resulted in higher achievement than the use of positive instances (examples) alone.

Questions: What is the correspondence between the definitions of "concept" and "strategy" used in these studies and in the Henderson studies? What, if any, effect does the focus on learning rather than on teaching have? Does the comparatively nonverbal presentation used in the CAI studies affect achievement?

A Comparative Comment

Dossey (in this monograph) summarizes a study he conducted which involved the use of examples and nonexamples in teaching disjunctive concepts (see pages 68-74 of this monograph). The example approach, in which the mix was one nonexample to two examples, was more effective in promoting student acquisition of disjunctive concepts than the nonexample approach, in which the mix was two nonexamples to one example.

Dossey's concern is obviously different from Shumway's. Dossey might label his treatments "example" and nonexample," but he is actually concerned more with the question, "What is the 'best' number of examples or nonexamples to use?" He accepts the efficacy of using nonexamples. Shumway, on the other hand, is asking, "Is the use of nonexamples helpful to students?"

The two studies each carefully controlled this variable, but on other variables their design differs. In his original study, Shumway (1971) conducted his research in a classroom setting, with teacher-pupil interaction uncontrolled on dimensions other than the independent variable. In later studies, he used a relatively nonverbal approach on CAI, with mathematical sentences being the mode of query. Dossey controlled the teacher effect by using programmed instruction. Comparisons across the Shumway and Dossey studies could be interesting because of these varying modes, _if_ they had been attacking the same question.

A serious limitation of Dossey's study is that the disjunctive concepts were contrived. In the studies Shumway has conducted on CAI, a parallel limitation arises, as students react to content which may appear similar to known content but actually is different. The transfer to reality is reasonable, but exploration with "real" content is also needed.

Comparisons can be interesting. Using what we learn from them, we can design studies that avoid previous problems. Sometimes we can even put results together and find verification for findings that is stronger because it comes from varying sources.

Comments on the Questions

The questions which have been raised in this section only begin to tap the reservoir of questions which could be asked. Some are very specific and can be answered by analysis or investigation. Some might provoke the design of a study. Some were intended to indicate concerns that do not at present appear to be reflected in the Henderson-oriented studies. In general, such factors as the organization and climate of the classroom, the awareness and expectations of the teacher, the type of training for interaction and questioning, the amount of student involvement, the time allotted, the age/grade level, the specificity of definitions, the type of language used, and the content selected for a study represent a few of the many facets that must be considered in planning research. A model for research on teaching strategies should make provision for them, especially so that continuing, coordinated research can be planned.

Deeper Reflections on One Study

Swank's paper (in this monograph) provides an opportunity to consider several specific facets of one study. He presents an example of a study that is carefully delineated and conducted which investigates (a) whether regulating the amount of content information contained in instructional strategies is a significant factor affecting student achievement and (b) the effect of teacher-pupil verbal interaction on student achievement. The points to be cited with reference to his study are ones which are important to consider as other studies are designed. Some of the questions in the previous section might also be applied to Swank's study (e.g., see questions for studies 1, 3, 6, 7, 9, 11, 13, 14, and 15).

Selection of Topic

Swank has met the primary criterion to be considered when evaluating research: Is the study attacking a question that is significant and worthwhile? His study is related to other Henderson-oriented studies, investigating a facet of the model not before studied, and in addition is related to a question of concern in teaching aside from its relationship to the model.

Specification of Variables

In Swank's study, as is frequently the case in research, the definitions of many variables are presented operationally. This is generally more than satisfactory for a given piece of research, and in some instances promotes replication and/or interpretation across studies. But there are problems in that the definitions do not always allow for the scope of reality. Consider, for instance, the "amount of content information" which is described in Swank's study by the number of content moves. Are all units of content information actually equivalent? Is it possible to make x number of concept moves under one strategy and y number under another strategy -- and thus affect the level of learning?

For the sake of the precision of operational definition, researchers must describe the variables within bounds -- but let us recognize that we are at the same time placing limitations on the findings. On the whole, however, the variables are clearly defined in this study, as are the hypotheses. The design is clear, incorporating many factors of concern.

Assumptions

As he investigates the effect of teacher-pupil verbal interaction on student achievement, Swank states an assumption attributed to Snow that "the probability that students are cognitively involved is directly proportional to the amount of overt participation." Are we then assuming that students do not learn from merely listening? Should we use verbal responses as the only measure of cognitive involvement? This question also pertains to studies other than this one; such assumptions must be logically analyzed.

Student Characteristics

The premise in this and other studies appears to be that the same type of interaction and the same content are suitable for average and high achievers. But might it be that the moves and strategies used in the study were actually more appropriate for the high achiever, and thus the results were influenced by this? Perhaps we need different strategies for different students. We verbalize the need for different approaches in teaching low achievers; we need to consider (and apply) far more understanding of individual differences as we design research and as we teach.

Control of Variables

Only one of many instances of careful control of variables in Swank's study will be cited: daily monitoring of the audiotapes. This assured that the treatment strategy was being implemented and is certainly a sound research technique. If the variables in a study are not well-controlled, the findings may be invalid or, at best, must be interpreted with care. Many studies are open to this criticism. Sometimes better control is unfeasible or impossible. At times, however, researchers overlook details, such as monitoring the lessons, that are quite possible.

Data Presentation

Swank's presentation and discussion of the data are carefully done; each finding is discussed and the conclusions are stated without overgeneralization. Presenting data is often well done in a research report. The discussion of the results and statements of conclusions are not so often done well: There is a great tendency to go further than the data warrant. When it comes to stating implications, some researchers manage to present everything they believe -- although these things may be far removed from the study being reported. This is particularly of concern since so often the relatively naive reader assumes that these are results from the research.

Swank's study, despite limitations, warrants replication and extension to other types of content and other types of knowledge. Subsequent studies related to Swank's work might consider the various questions identified above in the previous discussion.

Last Reflections

No attempt will be made to state conclusions or to summarize. Instead, five reflections will be restated -- points that must not be forgotten as we explore teaching strategies:

1. Teachers are individuals, and learn to teach in many ways. To analyze the process and to attempt to follow the analytic model may be the best way for some; modeling behaviors they see another teacher use may be the most effective way for others. Perhaps there are some teachers for whom specifying strategies is of more help than it is to other teachers. This factor must be considered as we select teachers for participation in research studies.

2. The reason for the focus on verbal behaviors is obvious -- but let us not completely overlook nonverbal communication. The facial expressions, the tone of voice, the rapport between teacher and learner all are important. Perhaps "rapport" is the ultimate ingredient of effective teaching.

3. Many things aid in making a teacher effective -- things like the teaching/learning environment, creativity, knowledge, ability to evaluate one's own behaviors, love of children, love of teaching -- in addition to command of strategies.

4. What may be an effective strategy for teaching mathematics at one level may not be appropriate at another level. Further, it has been observed that some moves are not possible with certain types of content. At this stage in the research process, let us be specific in stating what we believe may be the limits of generalizability for use of specific moves and strategies.

5. As we concentrate on teaching and the teacher, let us not forget: (a) the learner, (b) the context or environment, and (c) assumptions about each that must be made. For example, consider discipline and readiness. Teachers cannot teach effectively, no matter how well they can manipulate teaching strategies, if the learner does not want to or is not prepared to learn.

6. Let us continue to be systematic in the way we vary the different aspects related to strategies. Let us be careful in selecting and precise in defining the strategies being used, the content being taught, the environment in which the study is conducted (both school factors such as organizational structure and classroom factors such as the type of materials available), the attributes of the students being taught, the measures of student performance or learning. A description of the teacher aside from the treatment he is using and samples of the lessons and materials, rather than general descriptions which can be interpreted in varying ways, would also be helpful. Every research report should contain a concise but comprehensive description of any of the aspects of the research situation which might be pertinent.

References

Adhikary, P. K. Instructional effect on question asking behavior of prospective Peace Corps science and mathematics volunteer teachers (Doctoral dissertation, Indiana University, 1972). Dissertation Abstracts International, 1973, 33A, 4993. (University Microfilms No. 73-6962)

Bellack, A., Kliebard, A., Hyman, R. T., & Smith, F. L. The language of the classroom. New York: Teachers College Press, 1966.

Brown, J. K., Jr. Textbook use by teachers and students of geometry and second-year algebra (Doctoral dissertation, University of Illinois at Urbana-Champaign, 1973). Dissertation Abstracts International, 1974, 34A, 5795-5796. (University Microfilms No. 74-5534)

Cooney, T. J. An analysis of teachers' verbal behavior using the theory of relations (Doctoral dissertation, University of Illinois, 1969). Dissertation Abstracts International, 1970, 31A, 673. (University Microfilms No. 70-13,282)

Dimeolo, J. R. An experimental study relating elementary classroom teaching behavior to student achievement (Doctoral dissertation, Case Western Reserve University, 1968). Dissertation Abstracts 1969, 29A, 4219-4220. (University Microfilms No. 69-9338)

Dossey, J. A. The relative effectiveness of four strategies for teaching disjunctive concepts in mathematics (Doctoral dissertation, University of Illinois at Urbana-Champaign, 1971). Dissertation Abstracts International, 1972, 32A, 5467. (University Microfilms No. 72-12,142)

Fey, J. T. Patterns of verbal communication in mathematics classes (Doctoral dissertation, Columbia University, 1968). Dissertation Abstracts, 1969, 29A, 3040. (University Microfilms No. 69-3063) (a)

Fey, J. T. Classroom teaching of mathematics. Review of Educational Research, 1969, 39, 535-551. (b)

Flake, J. L. The use of interactive computer simulations for sensitizing mathematics methods students to questioning behaviors (Doctoral dissertation, University of Illinois at Urbana-Champaign, 1973). Dissertation Abstracts International, 1974, 34A, 7623. (University Microfilms No. 74-12,015)

Friedman, M. L. The development and use of a system to analyze geometry teachers' questions (Doctoral dissertation, Columbia University, 1972). *Dissertation Abstracts International*, 1973, *33A*, 4215-4216.

Gaston, J. A., & Kolb, J. R. A comparison of three strategies for teaching a selected mathematical concept to students in college algebra. *Journal for Research in Mathematics Education*, 1973, *4*, 177-186.

Gordis, F. W. A Piagetian analysis of the teaching of seriation concepts in four first grade classrooms (Doctoral dissertation, Columbia University, 1970). *Dissertation Abstracts International*, 1973, *33B*, 4994-4995. (University Microfilms No. 73-8949)

Gregory, J. W. The impact of the verbal environment in mathematics classrooms on seventh grade students' logical abilities (Doctoral dissertation, The Ohio State University, 1972). *Dissertation Abstracts International*, 1972, *33A*, 1585. (University Microfilms No. 72-27,014)

Heimer, R. T., & Lottes, J. J. Toward a theory of sequencing: An integrated program of research: The theoretical model and a synopsis of the first two years of the research program. *Journal for Research in Mathematics Education*, 1973, *4*, 85-93.

Henderson, K. B. A model for teaching mathematical concepts. *The Mathematics Teacher*, 1967, *60*, 573-577.

Hudgins, B. B., & Loftis, L. The invisible child in the arithmetic class: A study of teacher-pupil interaction. *Journal of Genetic Psychology*, 1966, *108*, 143-152.

Kester, S. W. The communication of teacher expectations and their effects on the achievement and attitudes of secondary school pupils (Doctoral dissertation, The University of Oklahoma, 1969). *Dissertation Abstracts International*, 1969, 30A, 1434-1435. (University Microfilms No. 69-17,653)

Kysilka, M. L. The verbal teaching behaviors of mathematics and social studies teachers in eighth and eleventh grades (Doctoral dissertion, University of Texas at Austin, 1969). *Dissertation Abstracts International*, 1970, *30A*, 2725. (University Microfilms No. 69-21,843)

Mahan, E. A. A study of the verbal behavior of four student teachers teaching selected geometric concepts to kindergarten children (Doctoral dissertation, Columbia University, 1970). *Dissertation Abstracts International*, 1971, *31A*, 3392-3393. (University Microfilms No. 70-26,791)

Meckes, R. C. A study to ascertain the instructional index and questioning strategy of mathematics teachers in grade 6, and to determine their relationship to professional characteristics and situational factors (Doctoral dissertation, Southern Illinois University, 1971). *Dissertation Abstracts International*, 1972, *32A*, 4245-4246. (University Microfilms No. 72-5380)

d

Merrill, M. D., & Wood, N. D. Instructional strategies: A preliminary taxonomy. Columbus, Ohio: ERIC Information Analysis Center for Science, Mathematics, and Environmental Education, 1974.

Nisbet, J. A. Instructional sequence for improving teacher question-asking in secondary mathematics (Doctoral dissertation, Arizona State University, 1974). Dissertation Abstracts International, 1974, 35A, 2841. (University Microfilms No. 74-25,806)

Rector, R. E., & Henderson, K. B. The relative effectiveness of four strategies for teaching mathematics concepts. Journal for Research in Mathematics Education, 1970, 1, 69-75.

Robitaille, D. F. Selected behaviors and attributes of effective mathematics teachers (Doctoral dissertation, The Ohio State University, 1969). Dissertation Abstracts International, 1969, 30A, 1472-1473. (University Microfilms No. 69-15,955)

Rollins, J. H. A comparison of three stratagems for teaching mathematical concepts and generalizations by guided discovery (Doctoral dissertation, University of Illinois, 1966). Dissertation Abstracts, 1966, 27A, 711-712. (University Microfilms No. 66-7802)

Sheppard, A. N. Changing learner conceptual behavior through the selective use of positive and negative examples (Doctoral dissertation, Indiana University, 1971). Dissertation Abstracts International, 1972, 32A, 4496. (University Microfilms No. 72-6726)

Shumway, R. J. Negative instances and mathematical concept formation: A preliminary study. Journal for Research in Mathematics Education, 1971, 2, 218-227.

Shumway, R. J. Human concept formation: Negative instances. Final report, NIMH Grant No. MH-20542, 1972.

Shumway, R. J. Negative instances in mathematical concept acquisition: transfer effects between the concept of commutativity and associativity. Journal for Research in Mathematics Education, 1974, 5, 197-211.

Smith, C. W. The effect of group appointed goals on mathematics teaching behaviors of elementary student teachers (Doctoral dissertation, The Pennsylvania State University, 1970). Dissertation Abstracts International, 1971, 32A, 827-828. (University Microfilms No. 71-21,803)

Smith, B. O., & Meux, M. O. A study of the logic of teaching. Urbana: University of Illinois Press, 1970.

Stilwell, M. E. The development and analysis of a category system for systematic observation of teacher-pupil interaction during geometry problem-solving activity (Doctoral dissertation, Cornell University, 1967). Dissertation Abstracts, 1968, 28A, 3083. (University Microfilms No. 68-893)

Strickmeier, H. B., Jr. An analysis of verbal teaching behaviors in seventh grade mathematics classes grouped by ability (Doctoral dissertation, University of Texas at Austin, 1970). Dissertation Abstracts International, 1971, 31A, 3428. (University Microfilms No. 71-198)

Vadya, K. B. A study of the relationship between classroom social-emotional climate and group problem-solving behavior in elementary school children (Doctoral dissertation, The Pennsylvania State University, 1967). Dissertation Abstracts, 1968, 29A, 158. (University Microfilms No. 68-8758)

Wolfe, R. E. Strategies of justification used in the classroom by teachers of secondary school mathematics (Doctoral dissertation, University of Illinois, 1969). Dissertation Abstracts International, 1969, 30A, 1064-1065. (University Microfilms No. 69-15,422)

Participants

1. Joe Dan Austin
2. Wilson Banks
3. David Bradbard
4. George Bright
5. Thomas Cooney
6. Edward Davis
7. David Davison
8. John Dossey
9. Janice Flake
10. William Geeslin
11. David Gliessman
12. John Gregory
13. Larry Hatfield
14. Kenneth Henderson
15. James Hirstein
16. John Kolb
17. Kil Lee
18. James Lockwood
19. William McKillip
20. Alan Osborne
21. Sidney Rachlin
22. Kenneth Retzer
23. James Rollins
24. Garth Runion
25. Larry Sachs
26. Richard Shumway
27. B. Othanel Smith
28. Curtis Spikes
29. Ronald Steffani
30. Leslie Steffe
31. Marilyn Suydam
32. Marc Swadener
33. Earl Swank
34. Richard Turner
35. Richard Weimer
36. Arthur White